

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 15, 2003, 18:55:03 ; Search time 5026 Seconds  
(without alignments)  
10711.713 Million cell updates/sec  
Title: US-10-033-269-8  
Perfect score: 1316  
Sequence: 1 tcggcagagctttctcag.....tggcattgtattgttagg 1316

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl :

- 1: gb\_ba.\*
- 2: gb\_btg.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	625	47.5	1338	6	A94667	A94667 Sequence 5
2	625	47.5	1338	8	LJPANC	Y10252 L.japonicus
3	599.6	45.6	1347	6	A94671	A94671 Sequence 9
4	422.2	32.1	1188	8	AY085534	AY085534 Arabidops
5	354	26.9	1284	6	A94668	A94668 Sequence 6
6	354	26.9	1264	8	OSPANC	Y10253 O.sativa pa
7	256	19.5	78973	8	AC102242	AC102242 Arabidops
8	246.4	18.7	118998	2	AC130200	AC130200 Medicago
9	226	17.2	145290	8	AC096688	AC096688 Oryza sat
10	147.6	11.2	4283	1	TNE7446	AJ007446 Thermotog
11	135.6	10.3	19711	1	AE001768	AE001768 Thermotog
12	131.6	10.0	302050	1	BX321856	BX321856 Nitrosomo
13	118.8	9.0	9381	6	AR204111	AR204111 Sequence
14	118.8	9.0	11214	1	AE005189	AE005189 Escherich
15	118.8	9.0	281530	1	AP002550	AP002550 Escherich
16	118.4	9.0	24845	8	SPAC5H10	Z49811 S.pombe chr
17	113.8	8.6	300933	1	AE016791	AE016791 Pseudomon
18	110.8	8.4	300409	1	AE016755	AE016755 Escherich
19	108	8.2	300029	1	AE016952	AE016952 Enterococ
20	107.6	8.2	304680	1	AE017002	AE017002 Bacillus
21	105.6	8.0	945	1	AY142936	AY142936 Helicobaci
22	105.6	8.0	11662	1	AE007789	AE007789 Clostridi
23	101.8	7.7	10709	1	AE015532	AE015532 Shewanell
24	101.6	7.7	10880	1	AE011810	AE011810 Xanthomon
25	101.2	7.7	4589	1	ECOPANBCD	I17086 E. coli alp
26	101.2	7.7	10982	1	AE000122	AE000122 Escherich
27	101.2	7.7	10982	6	AX370251	AX370251 Sequence
28	101.2	7.7	301250	1	AP004830	AP004830 Staphyloc
29	100.2	7.6	290525	1	AE017029	AE017029 Bacillus
30	99.8	7.6	690	6	AX143131	AX143131 Sequence
31	99.8	7.6	4103	1	AF270140	AF270140 Staphyloc
32	99.8	7.6	4103	6	AX145458	AX145458 Sequence
33	99.8	7.6	300275	1	AE016751	AE016751 Staphyloc
34	99.6	7.6	13180	1	AE015050	AE015050 Shigella
35	99.6	7.6	290029	1	AE016978	AE016978 Shigella
36	99.6	7.6	297850	1	AP003137	AP003137 Staphyloc
37	99.6	7.6	341350	1	AP003365	AP003365 Staphyloc
38	98	7.4	864	6	AX620824	AX620824 Sequence
39	98	7.4	52695	1	AC074026	AC074026 Staphyloc
40	97.6	7.4	1244	8	FOX298881	AJ298881 Fusarium
41	96.6	7.3	11217	1	AE003876	AE003876 Xylella f
42	96.4	7.3	314150	1	CJ11168X1	AL139074 Campyloba
43	96.4	7.3	331801	1	NMA422491	AL162755 Neisseria
44	95	7.2	300029	1	AE012553	AE012553 Xylella f
45	93.4	7.1	10929	1	AE004886	AE004886 Pseudomon

ALIGNMENTS

RESULT 1

A94667 A94667 Sequence 5 from Patent WO942565. linear PAT 26-JAN-2000  
LOCUS  
DEFINITION  
ACCESSION A94667  
VERSION A94667.1 GI:6778943  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified  
REFERENCE 1 (bases 1 to 1338)  
AUTHORS Laber,B. and Genschel, U.  
TITLE HERBICIDES TEST METHOD  
JOURNAL Patent: WO 942565-A 5 26-AUG-1999;  
LABER BERND (DE); GENSCHER ULRICH (DE)

QY	1082	GACACATGGAAATCAACTTGTCAATGAATCT----	TTGATCTAACCTTTG--TCATCT	1133
Db	923	GACACATAGAAATCAACTTGTAAATGGAATGATCTAACCTTGTGAATAATCT		982
QY	1136	CAAAATGGCCACATGCTTAATTAATAGTTTCGGGCCACGCTTAACAATTC		1188
Db	983	CAGACATGACCATATGATTAGTAGTCTTGGCATTTTCATGGGTATAGACTTC		1035
RESULT 2				
LJPANC				
LOCUS				
DEFINITION				
L.japonicus panc gene.				
ACCESSION				
Y10252				
VERSION				
Y10252.1				
GI:2292920				
KEYWORDS				
panC gene; pantoate-beta-alanine ligase.				
SOURCE				
Lotus japonicus				
ORGANISM				
Lotus japonicus				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;				
Lotus.				
REFERENCE				
1				
Genschel,U., Powell,C.A., Abell,C. and Smith,A.G.				
The final step of pantothenate biosynthesis in higher plants:				
cloning and characterization of pantothenate synthetase from Lotus				
japonicus and Oryza sativum (rice)				
Biochem. J. 341 (Pt 3), 669-678 (1999)				
JOURNAL				
MEDLINE				
FUBMED				
99348031				
REFERENCE				
2 (bases 1 to 1338)				
AUTHORS				
Genschel,U.				
TITLE				
Direct Submission				
JOURNAL				
Submitted (23-DEC-1996) U. Genschel, University Of Cambridge,				
Department Of Plant Sciences, Downing Street, Cambridge, CB2 3EA,				
UK				
FEATURES				
source				
Location/Qualifiers				
1..1338				
/organism="Lotus japonicus"				
/mol_type="mRNA"				
/sub_species="(Regel) Larsen"				
/db_xref="taxon:34305"				
/tissue_type="root nodules"				
/dev_stage="3 weeks"				
gene				
1..1338				
/gene="panC"				
CDS				
20..946				
/gene="panC"				
/EC number="6.3.2.1"				
/codon_start=1				
/evidence=experimental				
/product="pantoate--beta-alanine ligase"				
/protein_id="CAA71302.1"				
/db_xref="GI:2292921"				
/db_xref="SPTREMBL:O24035"				
/translation="MAPMTISDKDEMVKSRSMRSQGLIALVPTMGFLHEGLHLVLR				
DAHNADLVAISIVNPGPSPEDLSAPDSFGDLQKLMSVPGVDVVVPHPHNLVD				
YKGGDVAEACGGDGVSCVDRRSFGHETVWRABKLEKPLCGKRPVFRGVATIV				
TKLFNVPDVAVFGKKDYQWKIIQRNVRDLDFSIKVIGSEVIREKDGAMSRNVY				
LSPEREKAVSINKSLPRAKSAADGQIHCEKLNILVWQSIIEAGGRIDVAEIVDQNN				
LEKVEWKGPVVCVSAWFGKARLINIBINL"				
BASE COUNT				
380 a 242 c 328 g 388 t				
ORIGIN				
Query Match 47.5%; Score 625; DB 8; Length 1338;				
Best Local Similarity 79.0%; Pred. No. 1.7e-110;				
Matches 800; Conservative 0; Mismatches 195; Indels 18; Gaps 4;				
QY	194	GCCTCAAGGGTGATCTCCGACAAAGCCCTCGATGGAGAGCTGGTCGCGCTCGATCGGGCC	253	
Db	23	GCACCAATGGTGATATCTGATAAGACGAGATGCGAAATGGTCAAGTCCATCGATCC	82	
QY	254	CAGGCAAGCTCATTGGGCTGGTCCCAACCATGGGCTTCCCTCCAGCGGGCCACCTCTCG	313	
Db	83	CAAGGCAAGCTCATCGCCCTCGTTCACCATGGGCTTCCCTCCAGGAGGCCACCTTTCT	142	
QY	314	CTCGTGGCCAGCCCGCAACTCTCCAGCTGCTGCGCGCTCTCCATCTAGCTCAACCG	373	
Db	143	CTCGTCAGAGCGCTCACAACACAGCTGACCTCGTGGCGCTCTCAATCTATGTCAACCT	202	
QY	374	GGCCAAATTCGCCCCCAGGAGGACCTCTCCACCTACCCTCGACTTCGACGGGAGCTA	433	
Db	203	GGCCAGTTTTCGCCGACCGAGGACCTTTCGGCATACCTTCTGATTTCAAGGTGATCTC	262	
QY	434	AAGAACTCGGCTCGTTCCCGGGCGCTGACGCTGCTGTTTCCATCCCGTAACCTGTAC	493	
Db	263	CAAAAACCTCATGCTGTTCTCTGGTGTGTGATGTTGTTTCCACCCCAAAATTGTAT	322	
QY	494	GATTACGGGAAGACGGTGGTGTGACGTGGCAGAG-----GCTGTGGAATGGTGTG	547	
Db	323	GATTACGGTGTGATGGCGGTGATGCTGTGGCGAGTGTGGTGTGATGGGGTGGTGTCT	382	
QY	548	TGCGTTGA-----GAGTGGGTCCGGGCAAGAACTTGGGTGAGGTTGAGAAGCTGGAA	601	
Db	383	TGTGTTGATAGGAGGAGTGGTTTGGGCATGAACCTTGGGTAGAGCTGAGAAGCTGGAG	442	
QY	602	TTGGGCTGTGTGGGAAGACAGGCCCGCTTTCTTCAGAGGGGTGGCAGCTGTGGTGAAG	661	
Db	443	AAACCCCTTTGTGGGAAGATAGGCTGTTTCTTTAGAGGGGTGCCACCATTTGTACC	502	
QY	662	AAGTTGTTTAAATTTGTGAGCGCAGATGTGGCTGTGTTTCGGTAAAGAGATATCAGCAG	721	
Db	503	AAGTTGTTTAAATTTGTGAGCGCTGATGTTGCTGTGTTTGGGAAGAAGGACTATCAGCA	562	
QY	722	TGGCGCTTATTCAGAGGATGGTTCGAGATCTTGATTTTCCATAAAAGTGATAGGTGCT	781	
Db	563	TGGAATAATTATTCAGAGATGGTTCGAGATCTTGATTTTCCATAAAAGTGATAGGTGCT	622	
QY	782	GAATAACACGTGATTAATGATGGCTGGCAATGAGTTCACGTAATGTGCACCTTTCACT	841	
Db	623	GAAATTAACGTGAGAAAGATGGCCTAGCAATGAGTTCCTCGTAAATGTGTACCTACCT	682	
QY	842	GAAGAGAGGAAAGGCACTATCAATAATAATCATTTGTTAAGAGCTAAATCAGCAGCA	901	
Db	683	GAAGAGAGGAAAGGCGAGTATCTATAATAATAATCATTTTTAGAGCTAAATCAGCAGCA	742	
QY	902	GGAGATGGTCAGGTGCATTGTGAGAAGTTGACAAATTTGGTCATCCAAAGTGTACTGAT	961	
Db	743	GAAGATGGACAGATACATTTGAGAAATGATAAATCTTGGTCGTGCAAAAGTATCACCGAA	802	
QY	962	GCTGTGGGAAGGATCGATTATGCTGAGATTTGATCAAAATAATTTGGAGAAAGTGGA	1021	
Db	803	GCTGTGGGAAGGATTAATGATGCTGAGATTTGATCAAAATAATTTGGAGAAAGTGGA	862	
QY	1022	CAGATCAAGAGTCTCTGCTCTTCTGTGTTGCTGATGTTTGGCAAGTCAAGCTTATA	1081	
Db	863	TGGATCAAGGGTCTCTGTTCTCTGTTTCTGATGTTTGGGAAAGCCAGGCTTATA	922	

Db 83 CAAGGCAAGCTCATGCGCTCGTTCCACCATGGCTTCCTTCAGGAAGCCACCTTTCT 142  
Qy 314 CTCGTGCCCGCCGACCAACTCTCCGAGCTCGTCCCGCTCTCCATCTACGTCAACCCG 373  
Db 143 CTCGTGACAGAGCGCTCAACACCGCTGACTCGTCCCGCTCAATCTATGTCACCCCT 202  
Qy 374 GGGCAATTCGCCCCCAAGGAGACTCTCCACTACCCCTCCGACTTCGACGCGAGCTGA 433  
Db 203 GGGCAGTTTCCCGACCGAGGACCTTCCGCATACCCTTCTGATTTCAAGGTGATCTC 262  
Qy 434 AAGAACTCGGTCGCTCCCGGGCGGTGAGCTGCTTTTCCATCCCGCTAACTTGATC 493  
Db 263 CAAAACCTCATGCTCTCTCTGCGGTGATGATGTTGTTTCCACCCCAATTTGTAT 322  
Qy 494 GATTACCGGGAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547  
Db 323 GATTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382  
Qy 548 TCGCTTGA-----GAGTGGTCCGGCAAGAGTTGGGTGAGGGTTGAGAGCTGGAA 601  
Db 383 TGTGTTGATAGAGGAGTGTGTTTGGGCATGAAACTTGGGTAGAGCTGAGAGCTGGAG 442  
Qy 602 TTGGGGTGTGCGGAGAGAGCGCCCGTTTCTTCAGAGGGTGGCGACTGTGCTGACG 661  
Db 443 AAACCCCTTTTGGGAAGAGTAGCCCTGTTTCTTTAGAGGGGTGCCACATTTGATC 502  
Qy 662 AAGTTGTTTAAATTTGAGGACAGATGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721  
Db 503 AAGTTGTTTAAATTTGAGGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562  
Qy 722 TGGCGGCTTATCAGAGGATGGTTTCGAGATCTTGAATTTTCCATAAAGTGATAGTGCT 781  
Db 563 TGGAAAATTTATCAGAGATGGTTTCGAGATCTTGAATTTTCCATAAAGTGATAGTGCT 622  
Qy 782 GAATAACAGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841  
Db 623 GAAGTAATACGTGAGAAAGATGGCTAGCAATGAGTTCCCGTAAATGTTGATACCTT 682  
Qy 842 GAAGAGAGGAAAGGACCTATCAATAAATAAATCATTGTTAAGAGCTAATCAGCAGCA 901  
Db 683 GAAGAGAGGAAAGGACCTATCTAATAAATAAATCATTGTTAAGAGCTAATCAGCAGCA 742  
Qy 902 GGAGATGGTCAGGTGCTGAGAGTTCGAGAAATTTGACAAATTTGCTATCCAAAGTGTACTGAT 961  
Db 743 GAAGATGGACAGATACATTGTGAGAAATTCATAAACTTGTCTGTCGCAAGTATCACGAA 802  
Qy 962 GCTGGTGGAGGATCGATTATGCTGAGATTGTTGATCAAAATTAATTTGGAGAAAGTGAA 1021  
Db 803 GCTGGTGGAGGATCGATTATGCTGAGATTGTTGATCAAAATTAATTTGGAGAAAGTGAA 862  
Qy 1022 CAGATCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081  
Db 863 TGATCAAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922  
Qy 1082 GACAACATGGAATCAACTTGTCAATGAATGT---TTGATCTAACCTTCTG--TCATCT 1135  
Db 923 GACAACATGGAATCAACTTGTAAATGGAAGTGAAGTATGATCTAACCTTGTGAATATCT 982  
Qy 1136 CAACATGGGCCACATGCTTAATTAATAGTTTCGGGCACTGCTTTAACAATTC 1188  
Db 983 CAGACATGGACCATATGATTAGTAGTTCTTGGCATTTTCATGGGGTATAGACTTC 1035

RESULT 3  
A94671  
LOCUS A94671 1347 bp DNA linear PAT 26-JAN-2000  
DEFINITION Sequence 9 from Patent WO9942565.  
ACCESSION A94671  
VERSION A94671.1 GI:6778947  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified

unclassified.  
REFERENCE 1 (bases 1 to 1347)  
AUTHORS Laber, B. and Genschel, U.  
TITLE HERBICIDES TEST METHOD  
JOURNAL Patent: WO 9942565-A 9 26-AUG-1999;  
FEATURES LABER BREND (DE); GENSCHEL ULRICH (DE)  
Location/Qualifiers  
source 1..1347  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
exon 1..1347  
BASE COUNT 377 a 254 c 321 g 395 t  
ORIGIN  
Query Match 45.6%; Score 599.6; DB 6; Length 1347;  
Best Local Similarity 76.8%; Pred. No. 1.4e-105;  
Matches 805; Conservative 0; Mismatches 219; Indels 24; Gaps 5;  
Qy 252 CCAGGCGCAAGCTCATTTGGGTGCTGCCACCATGGGCTTCTCCACGGGGCCACTCT 311  
Db 69 CCCAAGCGAAGCTCATCGCCCTCGTTCCACCATGGGCTTCTTCAGBAGGCCACTTT 128  
Qy 312 CGTCTGTGGCCAGGCCCGCCAACTCTCCGACGCTGCTGCGGCTCCACTACGTCAACC 371  
Db 129 CTCTGCTCAGAGACGCTCAACACCGCTGACCTGCTGCGCGCTCTCAATCTATGTCAACC 188  
Qy 372 CGGGCCAAATTCGCCCCCAGGAGGACCTCTCCACCTACCCCTCCGACTTCGACGGGACG 431  
Db 189 CTGGCCAGTTTCCCGACCGGAGGACCTTTCGCATACCCCTCTGATTTTCAAGGTGATC 248  
Qy 432 TAAAGAACTCGCGTCCGTTCCCGGGCGGTGACGCTCGTTTCCATCCCGCTAACTGT 491  
Db 249 TCCAAAACCTCATGCTCTGTTCTGTGGTGGTGTGATGCTGTTTCCACCCCAATTTGT 308  
Qy 492 AGATTACGGGAAGAAGCGTGTGTGAGAGCTGGGAGAG-----GCTGTGGAATGGTGT 545  
Db 309 ATGATTACGCTGTGATGCGGCTGATGCTGTGCGGAGTGTGCTGTGCTGATGGGTGGTGT 368  
Qy 546 CGTGGCTTGA-----GAGTGGTCCGGCACGAAAGTTGGGTGAGGTTGAGAACTGG 599  
Db 369 CTTGTGTTGATGAGGAGTGGTTTGGGCATGAACTTGGGTGAGAGCTGGAAGCTGG 428  
Qy 600 AATTGGGGTGTGTGGAAGAGCAGGCCCGCTTCTTCAGAGGGGGTGGGCTGCTGCTGA 659  
Db 429 AGAAACCCCTTGTGGGAAGAGTAGSCCTGTTTCTTAGAGGGGTTCACCATTTGTTA 488  
Qy 660 CGAAGTGTGTTAATATTGTGGAGCCAGATGTGCTGTGTTCGGTAAAGAGGATTAATCAGC 719  
Db 489 CCAAGTTGTTAATATTGTGGAGCCTGATGTTGCTGTGTTGGGAAGAAGGACTATCAGC 548  
Qy 720 AGTGGCGCTTATTCAGAGGATGGTTTCGAGATCTTGAATTTTCCATAAAGTATAGGTG 779  
Db 549 AATGGAATAATTATTCAGAAATGGTTTCGAGATCTTGAATTTTCCATAAAGTATAGGTG 608  
Qy 780 CTGAATAACACGTGATAATGATGGCTCGCAATGAGTTTCACTAATGTGCACTTTTCAAC 839  
Db 609 CTGAAGTATACTGTGAGAAGATGGCTTAGCAATGAGTTCCCGTAAATGTTACCTATCAC 668  
Qy 840 CTGAAGAGAGGGAAGGCACTATCAATAAATAAATCAATTTGTTAAGAGCTAAATCAGCAG 899  
Db 669 CTGAAGAGAGGGAAGGCACTATCTATAAATAAATCAATTTTGTAGAGCTAAATCGGCAG 728  
Qy 900 CAGGAGATCGTCAGGTGCATTGTGAGAGTTGACAAATTTGGTTCATCCAAAGTGTACTG 959  
Db 729 CAGAAGATGGACAGATACATTGTGAGAAATTGATAAACTTGTGCTGCAAGTATCACCG 788  
Qy 960 ATGCTGCTGGAAGGATTCGATTATGCTGAGATTGTTGATCABAATAATTTGGAGAAGTGG 1019  
Db 789 AAGCTGCTGGAAGGATTTGATTATGCTGAGATTGTTGATCAAAATAAATTTGGAGAAAGTGG 848  
Qy 1020 AACAGATCAAGAGTCTGCTGCTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079





Qy	1025	ATCAAGAGTCTCTGTCGCTTCTCTGTTGCTGTCATCGTTTGGCAAGATCAAGCTTATAGAC	1084
Db	883	ATAAAGAGTGGGTAGTGATTTGTTGCTGCTTGGATTGGAAACGGTCAAGCTCATAGAC	942
Qy	1085	AAACATGAATCAACTTGTGCAATGAATGTTTGA	1117
Db	943	AACATTGAGATCAATGCTCTCTCTAGATCTCA	975
RESULT 5			
LOCUS	A94668	1264 bp	linear
DEFINITION	Sequence 6 from Patent WO9942565.		
ACCESSION	A94668		
VERSION	A94668.1	GI:6778944	
KEYWORDS			
SOURCE	unidentified		
ORGANISM	unidentified		
REFERENCE	1 (bases 1 to 1264)		
AUTHORS	Laber,B. and Genschel,U.		
TITLE	HERBICIDES TEST METHOD		
JOURNAL	Patent: WO 9942565-A 6 26-AUG-1999;		
FEATURES	LABER BERND (DE); GENSCHEL ULRICH (DE)		
source	Location/Qualifiers		
	1..1264		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
exon	/db_xref="taxon:32644"		
BASE COUNT	279 a 349 c 359 g	277 t	
ORIGIN			
Query Match	26.9%;	Score 354;	DB 6;
Best Local Similarity	63.2%;	Pred. No. 3.7e-58;	
Matches	584;	Conservative	0; Mismatches 325; Indels 15; Gaps 2;
Qy	187	GGCTCCAGCCCCAAGGGTGATCTCCGACAGGCGCTCGATCGGAGCTGTGTCGGCTCGAT	246
Db	83	GCGCGCGAGCCCGAGGTGATTCGCGCAAGGGCGGCGATCGCGCATGTGTCGGCGCGCG	142
Qy	247	GCGGGGCCAGGGCAAGCTCATTTGGCTGTGTCCCCACCACATGGGCTTCTCTCCACGCGGGCCA	306
Db	143	TCGCGCGAGGGCAAGACCGTTCGGGTCTGTACCCACCATGSGGTACTCTCCACGAGCCA	202
Qy	307	CTCTCTGCT-----CGTGGCCACAGGCCCGCCAACTCTCCGAAGTCTGTCGCCGT	354
Db	203	CCTCTCCCTCATCTCCGCCCGCCGCGCGCGCTCTCGCTGATCCCGTCGCCATCGTCGT	262
Qy	355	CTCATCTAGTCAACCCGGGCCAAATTCGCCGCCACGGAGAGCTCTCCACCTACCCCTC	414
Db	263	CACATCTAGTCAACCCGAGCAGTTCGGCGCCCTCAGAGGACCTCGCACCTACCCCTC	322
Qy	415	CGACTTCGACGGGCAAGCTAAAGAAACTCGCGTCCGTTCCCGCGGGCGGTCCAGCTCGTTT	474
Db	323	CGACTTCGCGGTGACCTTCGCGAAGCTCGACTCC---ACCGGCGTCTGTAATGCCGTCTT	379
Qy	475	CCATCCCGGTAACTTGTACGATTACGGGAAGAACGGTGTGTGTGAAGCTGGCAGAGCGTGG	534
Db	380	CAACCCCGCTGACCTCTACCTCGGTGCGCGCGCGTCCGCGGGCGCGGTCTCCGAGGCGC	439
Qy	535	TGGAATGGTCTCTGTCGTTTCAGAGTGGGTCCGGGACAGAAAGTTGGTGTAGGGTTGAGAA	594
Db	440	GATCTCTGCTCGAGGAGCGCGCGGGATGGGCACGACGTTGGTTCGGGTGGAGCG	499
Qy	595	GCTGGAAATTCGGGCTGTGTGGGAAGACGAGCCCGCTTTTCTTCAGAGGGGTGCGCACTGT	654
Db	500	ATTGGGAAGGGATTGTGCGGGCCAGCGCTCCGCTTCTTCGAGGGCGTGGCCACCAT	559
Qy	655	GGTGAAGAAGTTGTTTAATTGTTGGAGCCAGATGTGGCTGTCTCGTGAAGAAGATTGA	714
Db	560	AGTCTCCAAGCTGTTAACTATCATCGAGCGGGATGTTCTCTGTCTTCGGGAAGAAGATTGA	619

[illegible]

BASE COUNT	279 a	349 c	359 g	277 t
ORIGIN				

Query Match 26.9%; Score 354; DB 8; Length 1264;  
Best Local Similarity 63.2%; Pred. No. 3.7e-58;  
Matches 584; Conservative 0; Mismatches 325; Indels 1

Qy	187	GGCTCCAGACCCCAAGAGGTGATCTCCGACAAGGCTCTGATCGGAGCTGGTCGCGCTCGAT	246
Db	83	GCOCGCGAGCGGAGGTGATCCGCGACAAGAGCGCGATGCGGCATGTTGCGCGCCGCG	142
Qy	247	CGCGGCCACAGGCAAGCTCATTTGGCTGGTCCCCACCATGGCTTCTTCCACGCGGGCCA	306
Db	143	TCGGCCGAGGGCAAGACCGTCGCGGTCTGTACCCACCATGGGCTACTCTCACCAAGGCCA	202
Qy	307	CTCTCTCGCT-----CGTGGGCCAGGCCCGCCAACTCTCCGACGTGTCGCGGT	354
Db	203	CTCTCTCGCTCATCTCCGCGCGCGCGCGCGCTCGCTGATCCCGTCGCACTCGCTCGT	262
Qy	355	CTCCATCTAGTCTCAACCGGSCCAATTGCGCCCCACGAGGACCTCTCCACCTACCCCTC	414
Db	263	CACCATCTAGCTCAACCCAGCCAGTTCGCGCCCTCAGAGACCTCGCACCTACCCCTTC	322
Qy	415	CGACTTCGACGGCGACGTAAAGAAACTCGGCTCGCTTCCCGCGGCGCTGCACGTCGTTTT	474
Db	323	CGACTTCGCGGTGACCTCCGCAAGCTCGCCTCC--ACCGGCGCTCGTGGATGCGCTCT	379
Qy	475	CCATCCCGTAACTTGATACGATTACGGGAGAAAGGTGGTGGTCACTGGCAGAGGTGG	534
Db	380	CAACCCCCCTGACCTCTACGTCGTTGGCGCGCTGCGCGGGCGCGCTCCGAGGCGC	439
Qy	535	TGGAATGGTGTGTCGCTGAGAGTGGGTCCGGGCACGAAAGTTGGGTAGGGTTGAGAA	594
Db	440	GATCTCCTCGCTGGAGGAGCGCGGGATGGCACGAGACGTGGTTCGGGTGAGCG	499
Qy	595	GCTGGAAATGGGCTGTGTGGGAAGAGAGGCCCGCTTTCTTCAGAGGGTGGCAGCTGT	654
Db	500	ATTGGAAAGGGATGTGCGGGGCGACCGCTCCGCTGTTCTTCGAGCGCTGGCCACCAT	559
Qy	655	GGTGACGAAGTTGTTTAATATGTGGAGCCAGATGTGCTGTGTTTCGTAAAGAGGATTA	714
Db	560	AGTCCCAAGCTGTTTAACTATCATCGAGCCGATGTTCCTGTGTTCCGGAAGAAGATTA	619
Qy	715	TCAGCAGTGGCGCTTATTCAGAGATGGTTTCGAGATCTTGATTTTTCCATAAAGTGAT	774
Db	620	TCAGCAGTGGCGCTCATCTTGGCGCTATTTGGTCGGGACTTGATTTTGGCATAGATAAT	679
Qy	775	AGGTGCTGAATAACACGTGATATGATGGCTCGCAATGAGTTCAGTAAATGTGCACCT	834
Db	680	GGGATCAAGAAATTTGTCGGAACTGATGCTTTGCCATGAACCTCCCGAATGTGCACCT	739
Qy	835	TTACCTGAAGAGAGGGAAGGCACTATCAATAAATAAATTCATTGTTAAGAGCTAAATC	894
Db	740	ATCACGCGGAGGAGGGAAGGCAATTATCCATCAGTAGATCACTGTTGATGCTAGAAC	799
Qy	895	AGCAGCAGGAGATGGTCAGGTGCATTGTGAGAAGTTGACAAATTTGGTCATCCAAAGTGT	954
Db	800	TGGCGCCTTGAAAGGGAACACTGATTCAAAACAAATCAAAAACAAAATAGTACACACT	859
Qy	955	TACTGATGCTGGTGAAGGATCGAATTATGCTGAGATTTGTCATCAAAATAATTTGGAGAA	1014
Db	860	AACGTAAACTGCGGTCAGGTGACTATGTTGAGATCGTGGAGCAAGAAAGTTTGGTCCC	919
Qy	1015	AGTGGAACAGATCAAGAGTCTGTGCTCTTCCTGTGTTTGTCTCATGTTGGCAAGTCAAG	1074
Db	920	TGTAGAACAGATCGACGCCCCTGGGTCAATTGGCTTTCGGCGGTGTTTGAAGAGTTCAG	979

Qy	1075	GCTTATACACACATCGAATCAA	1098
Db	980	GCTGATCGATAATATCGAAATCGA	1003
RESULT 7			
AB012242/c			
LOCUS	AB012242	79973 bp	DNA linear
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K24G6.		
ACCESSION	AB012242	BA000015	
VERSION	AB012242.1	GI:3046850	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)		
AUTHORS	Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Miyajima,N. and Tabata,S.		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. VI. Sequence features of the regions of 1,367,185 bp covered by 19 physically assigned P1 and TAC clones		
JOURNAL	DNA Res. 5 (3), 203-216 (1998)		
MEDLINE	98403884		
PUBMED	9734815		
REFERENCE	2 (bases 1 to 78973)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-MAR-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:yinakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)		
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see <a href="http://www.kazusa.or.jp/kaos/cgi-bin/sgd_graph.cgi?c=K24G6">http://www.kazusa.or.jp/kaos/cgi-bin/sgd_graph.cgi?c=K24G6</a> Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <a href="http://compbio.ornl.gov/Grail-1.3/">http://compbio.ornl.gov/Grail-1.3/</a> ), GENSCAN (Chris Burge, MIT, <a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a> ), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <a href="http://www.cbs.dtu.dk/services/NetGene2/">http://www.cbs.dtu.dk/services/NetGene2/</a> ) and SplicePredictor (Volker Brendel, Stanford University, <a href="http://gremml1.zool.iastate.edu/cgi-bin/sp.cgi">http://gremml1.zool.iastate.edu/cgi-bin/sp.cgi</a> ). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <a href="http://genome.wustl.edu/eddy/tRNAscan-SE/">http://genome.wustl.edu/eddy/tRNAscan-SE/</a> ). This sequence may not be the entire insert of the clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K15N18 and the 3' clone is K19E20.		
FEATURES	Location/Qualifiers		
source	1. .78973 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="K24G6" /clone_lib="Miteui TAC" complement(1..49) /note="CDS is reported in Acc# AB015468 gb AAF20229.1 gene_id:K15N18.18 similar to unknown protein" /number=5 /evidence=not experimental complement(144..215) /note="CDS is reported in Acc# AB015468 gb AAF20229.1 gene_id:K15N18.18		
exon			
exon			





DEFINITION	Oryza sativa chromosome 3 BAC OSJNBa0015N08 genomic sequence, complete sequence.
ACCESSION	AC096688
VERSION	AC096688.4 GI:27228823
KEYWORDS	HTG.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1 (bases 1 to 145290) Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M., Overton, II, L.L., Tsitrin, T., Kim, M.M., Bera, J.J., Jin, S.S., Fedorosh, D.W., Tallon, L.J., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S.S., Riedmuller, S.B., Utterback, T.T., Feldblyum, T.V., Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.J., White, O., Salzberg, S.L. and Fraser, C.M.
TITLE	Oryza sativa chromosome 3 BAC OSJNBa0015N08 genomic sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 145290) Buell, R.
AUTHORS	Direct Submission
TITLE	Submitted (22-SEP-2001) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	3 (bases 1 to 145290) Buell, R.
AUTHORS	Direct Submission
TITLE	Submitted (18-DEC-2002) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA
REFERENCE	4 (bases 1 to 145290) Buell, R.
AUTHORS	Direct Submission
TITLE	Submitted (10-JAN-2003) The Institute for Genomic Research, 9712
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA, rbueller@igr.org
COMMENT	On Dec 18, 2002 this sequence version replaced gi:18092992. Address all correspondence to: rice@tigr.org
FEATURES	BAC clone OSJNBa0015N08 is from Oryza sativa chromosome 3 The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh ( <a href="http://www.softberry.com/">http://www.softberry.com/</a> ), Genscan and Genscan+ (Chris Burge, <a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a> ), GeneMarkHM (Mark Borodovsky, <a href="http://genemark.biology.gatech.edu/GeneMark/">http://genemark.biology.gatech.edu/GeneMark/</a> ), and GeneSplicer (Mihaela Pertea and Steven Salzberg, <a href="http://contact.mperetea@tigr.org">contact mperetea@tigr.org</a> ), searches of the complete sequence against a peptide database and the plant EST database at TIGR ( <a href="http://www.tigr.org/tdb/tgi.shtml">http://www.tigr.org/tdb/tgi.shtml</a> ). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, <a href="http://genome.wustl.edu/eddy/tRNAscan-SE/">http://genome.wustl.edu/eddy/tRNAscan-SE/</a> ). Simple repeats are identified by repeatmasker (Arian Smit, <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a> ).
SOURCE	This BAC overlaps with rice BAC OSJNBb0096M04 (AC092559) and OSJNBb006008 (AC120506). Location/Qualifiers 1. .145290 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="3" /map="near 1375" /clone="OSJNBa0015N08" /note="japonica cultivar-group" complement (116. .1322)
gene	<pre> /gene="OSJNBa0015N08.1" /note="EST C72591 from this gene" complement (join (&lt;116. .351,931. .1130,1204. .&gt;1322)) /gene="OSJNBa0015N08.1" complement (join (&lt;116. .351,931. .1130,1204. .1322)) /gene="OSJNBa0015N08.1" /codon_start=1 /product="unknown protein, 3'-partial" /protein_id="AAO20052.1" /db_xref="GI:27573334" /translation="MLNESMGEGDAAYAKREVLTTAGDDVSRGSIASLTATHGCRVLIV GREGALAGTAEEARCGGGGDAVAVGLDLHGCDCEAAVDAAGTAWRCDEGDLDMWNC YSEGEVQDCLNISEDEFKTKMKNVMTFWFLKATAKRLDRSESSCGGSLVLTQII GABRGLYPGAAAYGSLGAIHQIVR" 1399. .1431 /rpt_family="AT_rich" 1453_1515 /rpt_family="AT_rich" 1594. .1617 /rpt_family="TTTTc)n" complement (2130. .5799) /gene="OSJNBa0015N08.2" /note="contains Pfam profile PF03031 (NLI interacting factor)" complement (join (&lt;2130. .2324,3346. .3683,4248. .4676, 4905. .5001,5205. .5329,5649. .&gt;5799)) /gene="OSJNBa0015N08.2" complement (join (2130. .2324,3346. .3683,4248. .4676, 4905. .5001,5205. .5329,5649. .5799)) /gene="OSJNBa0015N08.2" /codon_start=1 /product="putative NLI interacting factor" /protein_id="AAO20054.1" /db_xref="GI:27573336" /translation="MPTKXKGLARNATTHINPNRKSRSRSTQAAAAHKVNDLITS SSKKIQGGLKKNRALKRGKKNLIYSAGTNDVAGVPSFFFNHKQSHENDERS CDSIFSPHNQKEDVDCLEGLVHEEVQNAEYVNTSI SCNTFDGLSKHSYD IMQACGSLTSGDEFSELGSPEVSAIYLAMQHSKLEICDSQDSTDCGADPD EYVELDYFDYAFIKDLPDLVWPQFVLLPKOTRSCPTTLVLDLDELTVHSTLE PCDSDFTPVFNREHTIVYRCPLYKEFETVASMEFEIIFITASQSIYAEQLLNI LDPKRLFRHYVRESCLFVEGNYLKDLVLGRLDARLVIVDNPQAFQGLDNGVFI ESWFDNRNDRELLTLFLPQSLVGVEDVRCPTARKFNLRKRVATAPSLSMHF" 6846. .6880 /rpt_family="AT_rich" 7034_10067 /gene="OSJNBa0015N08.3" /note="similar to cytochrome P450 CYP71E1 GB:AAC39318 GI:2766452 (Sorghum bicolor)" join (&lt;7034. .7534,8885. .9343,9444. .&gt;10067) /gene="OSJNBa0015N08.3" join (7034. .7534,8885. .9343,9444. .10067) /gene="OSJNBa0015N08.3" /codon_start=1 /product="putative cytochrome P450 protein" /protein_id="AAO20056.1" /db_xref="GI:27573338" /translation="MATLLSKLLALPQOQLLLLLLLLPIASLLIVIGENTGRRRRR HLRPPGPARLPVLGNLQLGALPHRSRLDARRHGPVNMRLRGVAVVSSPERAAQ EVLRTDADCCSRSPSPGFMRLUSYKQVAFAPYDAYGHAARLRFVAFELFSAPVQAA WRAQDQVEKLIKLTPEPEPEVELNDHIFALTDGII GAVAFSGIYGFERFAGGSKR FHHLLDDVMDLASFSAEDFFPNAARLFDHLGLVAHRERVFQOLDAFFBMVIEQH LDSNSAGGGGGLVGLIGLWKQKQYGDRRFRFTRNVKAIIFDAFIGIGTSSVTI LWMAELMESPRVMEKQVAEIRATYVGDGSGGVQPDLPRLAYLKVYVKEVLRLLHP ATLLMPRETRMDVRIGGYEVAARTVMVNAIGDAARWEEAEVDFDPRFAKRVFP NGHFELLPFGSGRRICPGIANGAANVEFTLANLLHCFDVALFVGAPFELSMESGG LVLHRKAPLVLPTRYIQL" 7082. .7106 /rpt_family=" (TCC)n" 8481_8595 /rpt_family="AT_rich" 10171. .10217 /rpt_family="AT_rich" complement (14707. .15477) </pre>

```
/gene="OSJNBa0015N08.4"
/note="predicted by fgenesh"
complement (<14707..>15477)
/gene="OSJNBa0015N08.4"
complement (14707..15477)
/gene="OSJNBa0015N08.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO20060.1"
/db_xref="GI:27573342"
/translation="MADLHPPEPTNGSSAAAAAALPDPAAEASGAQAAAG
PPYKRRRPSVRIGDTCVSAADADPLPQRHHKTSSSHPPPPRAHPEDALDPL
NHSSAAARRCQPAQRPRPAWPAAPGIDCYEEAEHYDDQSDSAAALARR
VSGRSGSDSDGVAGWLPNGRLPGSMGVGGVKGAWLDGLGLSRYPAPVEIHVDD
VLPLTLEDDKMDGIVAGVSRRLYAAIQLRSDSVS"
15138..15328
/rpt_family="G-rich"
repeat_region
15241..15357
/rpt_family="G-rich"
repeat_region
15395..15433
/rpt_family="(CGG)n"
repeat_region
17853..17883
/rpt_family="tandem"
repeat_region
19210..24290
/rpt_family="tandem"
repeat_region
/gene="OSJNBa0015N08.5"
/note="highly similar to aspartate kinase GB:CAA67376
GI:4376158 (Arabidopsis thaliana)"
Join(<19210..19593,19694..19776,19863..19903,21108..21262,
22007..22156,22277..22372,22629..22745,22860..22979,
23061..23153,23242..23358,23706..23822,23980..24042,
24129..24290)
/gene="OSJNBa0015N08.5"
Join(19210..19593,19694..19776,19863..19903,21108..21262,
22007..22156,22277..22372,22629..22745,22860..22979,
23061..23153,23242..23358,23706..23822,23980..24042,
24129..24290)
/gene="OSJNBa0015N08.5"
/codon_start=1
/product="putative aspartate kinase"
/protein_id="AAO20063.1"
/db_xref="GI:27573345"
/translation="MAVALRFAAVARDSPAAAPPRVGRQQVYLACAAAAPGRCR
ARGVLQCSGAAVLVKDDAAVAAAASSATGFTVAMKFGSSVAAERMEVAD
LIISFFPPTVVLSWAKGTNNLLAGEXACGPAKSEIPELAVIKELHVRIDE
LGLDRSVLSGLEQLKGVAMKELTRIDYLVSGECMASTRIFAAYLNKLGKK
ARQVADIDGFTTDDTADILEATPAVAKRLQGDWIDDPPIVITGLGKWKSC
AVTLGRGSDLTATTICKALGLREIOVKWDVGLTCDPNI CANAIPVPLPFEAA
ELAVFGAQLVHPQSMRPAKEDIPIRVKNSVNRAPCTVITKTRDLSKSLTSLVLS
NITWLDIVSTRMLQYGLAKVFSIPEDLGISVDCVATSEVSLTDPKWLKRELI
QQELDHVVELEKIAVHLHQHRSIISLIGNVQRSSILEKAFNVLRRNVNWMISQ
GASKVNISLVVHDSKQCIQALHSAFFENGFLUSEVEEDLQHNQSPVSPNGVIYGN"
21672..21726
/rpt_family="tandem"
repeat_region

Query Match 17.2%; Score 226; DB 8; Length 145290;
Best Local Similarity 62.9%; Pred. No. 3e-33;
Matches 390; Conservative 0; Mismatches 215; Indels 15; Gaps 2;

Qy 187 GGCTCCAGCCCCAAGGGTGATCTCCGCAAGCGCTCGATCGGAGCTGGTCGCGCTCGAT 246
Db 94344 GCGCGGAGCGGAGGTGATCCGCGACAGCGCGGATGCGCGCATGTGTCGCGCGCGG 94403

Qy 247 GCGGGCCAGCGCAAGCTCATTTGGCTGGTCCCAACCATGGGCTTCCTCCACGCGGGCCA 306
Db 94404 CCGCGCGAGGCGAAGCCGTCGCGGTGCTACCCACCATGGCTACCTCCACCAAGGCCA 94463

Qy 307 CCTCTCGCT-----CGTGCCAGCGCGGCAACTCTCCGACGTCGTCGCGGT 354
Db 94464 CCTCTCGCTCATCTCCGCGCGCGCGCGCGCTCCGCTGATCCGCTGCGCATCTGCTGT 94523

Qy 355 CTCCTACTAGCTCAACCGGCGCAATTCGCCGCCACGAGGACCTCTCCACCTACCCCTC 414
Db 94524 CACCATCTAGCTCAACCGGCGCAATTCGCCGCCCTCAGGAGACCTCGCCACCTACCCCTC 94583
```

```
Qy 415 CGACTTCGACGGCGACGTAAGAACATCGGTCGGTCCGTCGGGGCGGCTCGACGTCGTTTT 474
Db 94584 CGACTTCGCGCGTGACCTCCGCAAGCTCGCCTCC---ACCGGCGTCGTGATGCGCGCTT 94640

Qy 475 CCATCCCGCTAACTTGTACGATTACGGGAAGACGGTGGTGGTACGTCGCGAGGCTGG 534
Db 94641 CAACCCCGCTGACCTTACGTCGCGCGCGGTCGCGCGCGCGCTCCGAGGCGC 94700

Qy 535 TGGAAATGGTGTGTCGTCGCTTTCAGAGTGGTCCGCGGACGAAAGTTGGTGGAGGTTGAGAA 594
Db 94701 GATCTCTGCTGGAGGAGCGCGCGGGGATGGGACGACGATGGGTTGGGTTGGAGCG 94760

Qy 595 GCTGGAATGGGCTGTGTGGGAAGACGAGCCCGCTTTCTTTCAGAGGGGTGGCGACTGT 654
Db 94761 ATTGGAGAAGGATGTGCGGGCGCAGCGCTCCGCTGCTTCTCCGAGCGTGCCACCAT 94820

Qy 655 GGTGACGAAGTCTTTAAATATTGTGGAGCCAGATGTGGCTGTGTTGGTGAAGAGTTA 714
Db 94821 AGTCTCAAGCTGTTAAACATCATCGAGCCGATGTGCTGTGTTGCGGAAGAGGATTA 94880

Qy 715 TCAGCAGTGGCGGCTTATTCAGAGGATGGTTCGAGATCTTGATTTTCCATATAAAGTGAT 774
Db 94881 TCAGCAGTGGCGGCTATCTGCGGTATGGTAAACAAATCTGCCAATTTCTTCTTTT 94940

Qy 775 AGGTGCTGAATAACACGTG 794
Db 94941 ATCTTTCATTTACATATG 94960

RESULT 10
TNE7446
LOCUS
DEFINITION Thermotoga neapolitana top2A and lexA genes, and ORF3 to ORF7.
ACCESSION AJ007446
VERSION AJ007446.1 GI:3282088
KEYWORDS lexA gene; LexA repressor; subunit A; top2A gene; topoisomerase II.
SOURCE Thermotoga neapolitana
ORGANISM Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
REFERENCE 1
AUTHORS Zverlov,V.V. and Schwarz,W.H.
TITLE Organization of the chromosomal region containing the genes lexA and topA in Thermotoga neapolitana. Primary structure of LexA reveals phylogenetic relevance
JOURNAL Syst. Appl. Microbiol. 22 (2), 174-178 (1999)
MEDLINE 9319341
PUBMED 10390868
REFERENCE 2 (bases 1 to 4283)
AUTHORS Zverlov,V.V.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1998) Zverlov V.V., Russian Academy of Sciences, Institute of Molecular Genetics, Kurchatov Sq., 123182 Moscow, RUSSIA

FEATURES
source
1..4283
/organism="Thermotoga neapolitana"
/mol_type="genomic DNA"
/strain="22706-MC24"
/db_xref="taxon:2337"
1..421
/gene="top2A"
<1..421
/gene="top2A"
/EC number="5.99.1.3"
/codon_start=2
/transl_table=11
/product="topoisomerase II subunit A"
/protein_id="CAA07512.1"
/db_xref="GI:3282089"
/db_xref="SPTREMBL:O86947"
/translation="IIRFPVSDVRMRNAGVQAIKLOFGDEVMSDVVPBEEGIL
TVTEKFGKRTPVQLRYRVQKRGGTGLRNTSDVNKTHVAVRYVKGEDEEIVVYTRNGM
```



Eisen, J.A., Fraser, C.M. et al.  
Evidence for lateral gene transfer between Archaea and bacteria  
from genome sequence of *Thermotoga maritima*  
Nature 399 (6734), 323-329 (1999)

99287316  
PUBMED  
10360571  
2 (bases 1 to 19711)  
Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J.,  
Hart, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A.,  
McDonald, L., Ufferback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M.,  
Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A.,  
Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D.,  
White, O., Salzberg, S.L., Smith, H.O., Venter, J.C. and Fraser, C.M.  
Direct Submission  
Submitted (01-JUN-1999) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
Location/Qualifiers  
1. 19711  
/organism="Thermotoga maritima"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2336"  
complement (67. .978)  
/gene="TM1075"  
complement (67. .978)  
/gene="TM1075"  
/note="similar to percent identity: 0.00; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAD36152.1"  
/db\_xref="GI:4981620"  
/translation="MKDGLYVAFSGKSSLVAILAKMALGEERVELTVDWSPYTY  
ERREIVRNFAEKHLKHTIPSNRMOEKVKGHPSCNACTRDVKTIVLRYAGHLV  
ASGANADSGKTKGVDFGVYSPLCRVKEENEMKFLGLEKVKIGESAGREGCKL  
KHLKMLNPDIYHGKAVSTANEILLRVLEEGEPKELANUKIIGPLSRNIALVNVKPL  
PPEKVMNEIVEKLSAESTIDGVIIVDFGPKMLVLASPAIYRNESRKRKEGRQLPEF  
AFPIEWRSEKNNKRTFOVVDARKE"  
complement (1008. .1481)  
/gene="TM1076"  
complement (1008. .1481)  
/gene="TM1076"  
/note="similar to GB:AE000782 percent identity: 51.61;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAD36153.1"  
/db\_xref="GI:4981621"  
/translation="MKRFLISDSHVVRMASLPDEILNSLKEYDVGIVGLGDYVDLDT  
VILLESKEFYGVHGMYPDYKHLFPFSKVLIVGVTTCMGWGAPWDLKDLIK  
VFNEKQVILFGHTHEPDTVKAGVPLNPSLAEGSYAVLELDGGEVRPELKT"  
complement (1486. .2328)  
/gene="TM1077"  
complement (1486. .2328)  
/gene="TM1077"  
/note="similar to GB:AE000657 percent identity: 74.16;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="pantoate--beta-alanine ligase"  
/protein\_id="AAD36154.1"  
/db\_xref="GI:4981622"  
/translation="MRIETIEEMKFSERREKKTIGFVPTMGYLHEGHSILVRRRA  
RAENDVVVSVIFVNPOTFGNEDYRYPDRFDRKLLKENVDICIFHPSVEEYPPD  
FSYVETKLSKHLGSRSGCHRGCTVTKLVNVPKRAIFGQDQQRVLRMM  
VRDLNMDVMEICPIVREPDGLAWSRNVLSPERQALSQSLKIAENLYNGER  
DAEKIKEMIKHLISDFDKIDYVEIVDEETLEFVEKIDRKIVAVAAWGNARLDIN  
TILG"  
complement (2325. .2837)  
/gene="TM1078"  
complement (2325. .2837)  
/gene="TM1078"

/note="similar to GB:AE000657 percent identity: 65.74;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAD36155.1"  
/db\_xref="GI:4981623"  
/translation="MGDPDRDEKIVGMQLGRKITNMRVVFRCVPGYVVIETSPKIDG  
KPPTLYWTCPLHREKVSQGVKSEFEKLEKGLREKWKKAHVEILIKRAEL  
LPEDHPRLTILSVGTGGTRDFTKLKCLHLHLADYLAGVKNPVGWVWDLVERKCEE  
KLCRAGEARE"  
complement (2807. .3601)  
/gene="TM1079"  
complement (2807. .3601)  
/gene="TM1079"  
/note="similar to GB:AE000657 percent identity: 61.19;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein L11 methyltransferase,  
putative"  
/protein\_id="AAD36156.1"  
/db\_xref="GI:4981624"  
/translation="MRFKELILFLKIEEELVEKFEYEGFFNFAIEDKKGKVLKIY  
LRGEPLPDFKDWIWDKITTPKDWIWELEPFIVEGIFIDPTEKINRDAIVKL  
SPGVAFGTGLHPTTRMSVFLKYLKEGTVLDVGGTGILATAAKLGASRVAVDV  
DEQAVAEENVRKNDVDVLKWSDLLSEVEGTFDIVSNILAEIKLLEDVNRVTH  
ROSMILTSIGIVDERKEDMKRKAASEHGNVLERQEREWTVLVMKRS"  
complement (3591. .4022)  
/gene="TM1080"  
complement (3591. .4022)  
/gene="TM1080"  
/note="similar to GB:AE000657 percent identity: 77.62;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="sugar-phosphate isomerase"  
/protein\_id="AAD36157.1"  
/db\_xref="GI:4981625"  
/translation="MKIAISDHAFAELKEKVNKYLKGIEVEDHGTYSSESVDPD  
YAKVVOISLSNEADFGILCGTGLGMSIAANRYGIRALCLFPDMARLARSHNAN  
ILVLPORLIGAEAFWIVDTFLTFPDGGRHERRIKIDEV"  
complement (4019. .4360)  
/gene="TM1081"  
complement (4019. .4360)  
/gene="TM1081"  
/note="similar to GP:2660753 percent identity: 64.15;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="anti-sigma factor antagonist, putative"  
/protein\_id="AAD36158.1"  
/db\_xref="GI:4981626"  
/translation="MFPYKIVDDVILMPNKNELNIENAHLFKKWVDFELNKNYKIF  
LVLSDEVESIDFSLGVIVNLKSISSGGFFALVSPNEKVERVLSLTNLDRIVKIYDT  
ISEAMEVEVRK"  
complement (4372. .4965)  
/gene="TM1082"  
complement (4372. .4965)  
/gene="TM1082"  
/note="similar to GP:2645227 percent identity: 100.00;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="lexA repressor"  
/protein\_id="AAD36159.1"  
/db\_xref="GI:4981627"  
/translation="MKDLTERORKVLLFIEEFIEKNGYPPSVREIARPRITPRGALL  
HLIALEKGYIERKNGKPRALRISIRNKIPLEIGIRAGEKREAIYEDLTIPEPS  
FLSSGVDHPLLKVGESMIEEHICDGLVLVRRQDWAQNGDIVAAWVGEVTLKKFYQ  
RGDTVELPARNEMSSNFRAEKVKILGKVGWVFRKL"  
complement (4962. .5336)  
/gene="TM1083"



CDS	complement (4962..5336) /gene="TM1083" /note="similar to GB:L77117 SP:Q60334 PID:1498783 percent identity: 53.91; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAD36160.1" /db_xref="GI:4981628" /translation="MRKPIETHADIAVEYISGNSYEEELLEARNILLLEEIGVILDTBEK EKYPIETEDAFDFTVNDVILRISKGWAPRIKREGNELKVTFRKIRKEGTEIKAL TYHLKPERDGDVLKTKVFDT" complement (5333..7747) /gene="TM1084" complement (5333..7747) /gene="TM1084" /note="similar to GP:2645226 percent identity: 99.88; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="DNA gyrase, subunit A" /protein_id="AAD36161.1" /db_xref="GI:4981629" /translation="MPEILINKPVEDLBSVLLYSMSVIVGRAIPDVRDGLKPVQRRLIYMYELGLGKHNSPTKKSARIYGEVMKYHPHGDAPVYDALVRMAQPTMRVPLIEG QNFGSIRDRPPAMRYTEARLRLAEMLLEIEKNTVNDIMDFDGLKEPEVPLSKV PNLINGASGIAGVMAINIPPHNLSETVDALYLIIDHEATVEELMQFIKGPDPFKA VVYNASLEGLYBEGRIIVRGKHVDGKRVKRIIVITELPYGSVAGLIEQIATKA KDDSLP IIRIRDSKRGRIIVIEIPKDAEEBVIINLYKRTALQDYFNVMQVLDK KDFRLMLNGLMEAFLEHREFEVIRRRAYEQYTRAHVVEGLKARAIGVVVDI VRNSKVESARQSLMLETITEBOAKALLDRLSLTSLEIENLNEYSDLRKISEV KEILEKDKVKIMKEFLYLKQYGDPRTEVDQSIENEYNEBELIVEDVITLSHK GYDKSTPLNSYKQGGKGTIVKLSDEDEVEFVVAKNTSLFTITNLGRAYVLKN YQLETCRNTGRHTAFNLDETEKIVALASINGRDLIVATKSGIKRTALKKFE NATSNRGRAIKPEGDEIVSARVNSKETEILIVATKMGMAIRFPVSDVRMGRNAG VOAILQPGDEVVSVDVIRPGEIEILTVTEKGFKEITPVQLYRIQRGGTGLRNISD VNKTGYVAVRYVRGDBEIVVTRNGMIRFPVSEIGVIGRVTKGKLIELGDDTITSK VAVVKD" complement (7761..9650)	QY 1009 GGAGAAAGTCGAAACAGATCAAGAGTCCTGTCGCTCTTGTGTGTCGATGTTGGCAA 1068 Db 1579 GGAGCCGGTTGAGAAATCGATCGAAAGGTGATCGTCGCGTGGCTGGCTGGGTTGGAAA 1520 QY 1069 AGTCAGGCTTATAGACAAAGATG 1090 Db 1519 TGCAGACTCATAGATAACACG 1498
gene	RESULT 12	
CDS	LOCUS EX321856 302050 bp DNA linear BCT 23-APR-2003 DEFINITION Nitrosomonas europaea ATCC 19718, complete genome; segment 1/10. ACCESSION EX321856 AL954747 VERSION EX321856.1 GI:30179916 complete genome. KEYWORDS Nitrosomonas europaea ATCC 19718 Nitrosomonas europaea ATCC 19718 Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae; Nitrosomonas. REFERENCE 1 Chain, P., Lamerdin, J., Larimer, F., Regala, W., Land, M., Hauser, L., Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Archiero, D., Holmes, N., Whittaker, M. and Arp, D. Complete Genome Sequence of the Ammonia-Oxidizing Bacterium and Obligate Chemolithoautotroph Nitrosomonas europaea J. Bacteriol. 185 (9), 2759-2773 (2003) 12700255 2 (bases 1 to 302050) Larimer, F. Direct Submission Submitted (12-NOV-2002) Submitted on behalf of the Nitrosomonas genome consortium, the DOE Joint Genome Institute, Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA, and the Genome Analysis Group, Oak Ridge National Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA; larimerf@ornl.gov Nitrosomonas genome consortium Location/Qualifiers 1..302050 /organism="Nitrosomonas europaea ATCC 19718" /mol_type="genomic DNA" /strain="ATCC 19718" /db_xref="taxon:228410" 211..1590 /gene="dnaA" /locus_tag="NE0001" 211..1590 /gene="dnaA" /locus_tag="NE0001" /function="InterPro IPR001687:IPR001957:IPR003593 COG593" /codon_start=1 /evidence=not_experimental /transl_table=11 /product="dnaA; chromosomal replication initiator protein" /protein_id="CAD83912.1" /db_xref="GI:30179917" /translation="MQKIEFWHFCFLKHPRELNQFNTWKPKLVKVCPEKNTLI LIAPNRFVQWIKNFVTRIDEMAQDHENRISPRLEFEPAESQAVRTVSAQKNRE DKPAEKTCQGVTSKRTNPSQNASFTFADVTGKANQLARAGATQVAERPGIAYNPL FYGGVGLGKTHLMOAIGNYVLELDAGAKIRYVHAEKYVSDVSAVQHSKDPKFLYY HSLDLLLVDDVQFTSGKNTQBEFFYAFNALIEAHKQVITTSDCVPEKISGLEELVS RFGWGLTVAIPEPELMEVAILLKKALAEKIELDENTAFFIAKYIRSNVRELEGAKR VLASRFTHSISLDLAEKALLKALQNRQISIENTQKVADYIKIKVADWYSKRV RTIVRQVAVNAIAKEULQSLSDPDGEAFGRDHTTVLHAHRKIELFTSDPGINRDF NALMHILRG" 1791..2912 /gene="dnaN" /locus_tag="NE0002" 1791..2912 /gene="dnaN" /locus_tag="NE0002"	
gene	gene	
CDS	CDS	



```

QY 668 TTTTATATTTGGAGCCAGATGTGGCTGTGTTTCGGTAAAGAAGGATTAATCAGCAGTGGCGG 727
Db 92419 TTCAATATCGTGCAGCCCATATAGCTGTGTTGGTGAATAAGACTATCAGCAGTTGCAG 92478
QY 728 CTTATTTCAGAGAGTGGTTTCAGATCTTGAATTTTTCATAAAGATGATAGTGTCTGAATA 787
Db 92479 GTTGTTCAGAGATGGTTTGAATCACTGAACCTTACCGGTAGAGATCAATTCGAGGTGAACA 92538
QY 788 ACAGTGTATATGATGGCTGGCAATGAGTTTCAGTAATGTGCACCTTTACCTTCAGAGAG 847
Db 92539 GTGCGGATGAGGATGTCTGGCGCTCAGTTCCCGCATATTAACCTGATGCTACCCAG 92598
QY 848 AGGGAAGAGGCACTATCAATAATAATCATTTGTTAAGAGTAATATCAGCAGCAGGAGAT 907
Db 92599 CGCAGGAGGAGGCGAATTTGGCACACCATCTGAATAACAGATCAGAGACAGCATCGCTTCA 92658
QY 908 GGTCAAGTGCATTTGTGAGAAGTTGACAAATTTGGTCATCCAAAGTTTACTGATGCTGGT 967
Db 92659 GGTGAACGGGACTTTCCCGCTGTGGAACAACTTGCAGCGGAAATTTGAGCAAGCGTGGC 92718
QY 968 GGAAGGATCGATTATGCTGAGATTTTTCATCAAAATTAATTTGGAGAAGTGGAAACAGATC 1027
Db 92719 TGGGTGTCGATTATGTTGCGTGCCTCAGCAACATACGCTGCTGCCAGTGGCAGCAAGC 92778
QY 1028 AAGAGTCTGCTGCTTCTGTGTTGCTGTCATGGTTTGGCAAAAGTCAGGCTTTATAGACAAC 1087
Db 92779 GATTCCAGCCCTCGTTATCTAGGCGCAGCATGGCTGAATCAGACGCGACTGATCGACAAT 92838
QY 1088 ATGGAATCAACTTTGTCA 1105
Db 92839 TTCCTGCTGACGTGGCA 92856

RESULT 13
AR204111/c
LOCUS 9381 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 7 from patent US 6365723.
ACCESSION AR204111
VERSION AR204111.1 GI:21500673
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Blattner, F.R., Burland, V., Perna, N.T., Plunkett, G. and Welch, R.
TITLE Sequences of E. coli O157
JOURNAL Patent: US 6365723-A 7 02-APR-2002;
FEATURES
source
BASE COUNT 2363 a 2263 c 2218 g 2537 t
ORIGIN

Query Match 9.0%; Score 118.8; DB 6; Length 9381;
Best Local Similarity 48.3%; Pred. No. 1.2e-12;
Matches 427; Conservative 0; Mismatches 397; Indels 60; Gaps 1;

QY 239 CGCTCGATCGCGGCCCAAGCTCATTTGGGCTGCTCCCAACATGGGCTTCTCTCCAC 298
Db 1111 CGCGCCTGCTATGGAAGGCAAGCGGTGGCTGTGTAACCTACCATGGGCAACCTGCAT 1052
QY 299 GCGGGCCACCTCTCGCTGTTGGCCAGGCGCCGCAACTCTCCGACGCTGCGCGTCTCC 358
Db 1051 GATGGCCACATGAGCTGTGTCAGAGAGCAAGCCGCGCGGATGTGTCGTGCTCAGT 992
QY 359 ATCTACGTCACCGGGCCCAATTCCGCCCAAGGAGGACCTCTCCACTACCCCTCCGAC 418
Db 991 ATTTTCGTTAACCGGATGTCAGTTCGACCGCGCGGAGGACCTGGCTCGGTACCCAGCACC 932
QY 419 TTCGACCGGCGGTAAGAACTCGCGTCCGTTCCCGGGCGGCTGACGTCGCTGTTTCCAT 478
Db 931 TTCAGGAAGACTGCGGAAAAGCT----- 909

```

```

QY 479 CCCCGTAACCTTGTTACGATTACGGGAAGAACGGTGGTGGTACGTCAGTGGCAGAGGCTGGTGA 538
Db 908 -----GAAACAAGCCAAATGGGATCTGGTTTTCGCGCCCTCG 872
QY 539 ATGCTGTCGTCGTTGAGAGTGGGTCCGGGCAAGAAAGTTGGGTGAGGGTTGAGAAGCTG 598
Db 871 GTAAAAGAGATCTACCCGAACGGTACTGAAACCCATACCTTACGTTGACGTTCCAGGGCTT 812
QY 599 GAATTTGGGGCTGTGTTGGAAGAGCAGGCGCGTCTTCTTTCAGAGGGTGGCGACTGTGGTG 658
Db 811 TCGACCATGCTGGGAAGCGCGCAGCCGCGCGGCAATTTTCGCGCGCTGTGCGACTATCGTC 752
QY 659 ACGAAGTTGTTTAAATATTTGTGGAGCCAGATCTGTGCTGTGTTTCGCTGAAGAAGGATTATCAG 718
Db 751 AGCAAGCTGTTTAACTGTGTCAGCCGCAATCGCTGCTTCGGTGAAGAGACTTTTCAG 692
QY 719 CAGTGGCGCTTATTCAGAGATGGTTCGAGATCTTGAATTTTCCATAAAGTGAATAGGT 778
Db 691 CAACTGGCGCTGATCCGCAAAATGGTTGCTGATATGGCTTTGATATTGAGATTGTCGT 632
QY 779 GCTGAATACACGTCGATATGATGGCTGGCATGAGTTTCACTAATGTGCACCTTCA 838
Db 631 GTGCCAATTTATGCGCGCCAAAGACGGTCTGCGCACTGAGTTCCCGTAAACGGTTATCTCAG 572
QY 839 CCTGAAGAGAGGGAAGGCACTATCAATAAATAAATCATTTGTTTAAGAGCTAAATCAGCA 898
Db 571 GCAGAACACGCAAAATTTGCCCGGCTCTGTACAAAGTTTAAAGTTTCGATTGCCGACAA 512
QY 899 GAGGAGATGTCAGGTGCATTTGAGAAGTTGACAAATTTGGTTCATCCAAAGTGTACT 958
Db 511 TTGCAGGCTGGCGAAACGGGATCTCGATGAAATTTATTCCTATTCGCGGCAAGAACTGAAT 452
QY 959 GATGCTGCTGGAAGATCGATTATGCTGAGATGTTGATCAAAATAATTTGGAGAAAGTG 1018
Db 451 GAAAAGGCTTCCGCGCGATGATATTCAGATTCGCGATCCGACACATTTCTCTGGAAGTC 392
QY 1019 GAACAGATCAAGAGTCTCTGCTGCTTCTGTTGCTGTCATGGTTTGGCAAGTCAGGCTT 1078
Db 391 TCTGAAACACGACAAACGGGCAAGTAATTCCTGTAGCCGCTGGCTTGGGATGTCGCGCTG 332
QY 1079 ATAGACAACATGGAATCAACTTTGTCAATGAATGTTTGATCTAA 1122
Db 331 ATCGACAACAAATGGTCGAGCTGGCGTAATACTTATCTCTAA 288

RESULT 14
AE005189/c
LOCUS
DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 1 of 3, section 13 of 155.
ACCESSION AE005189
VERSION AE005174
KEYWORDS
ORGANISM Escherichia coli O157:H7 EDL933
REFERENCE
AUTHORS Escherichia coli O157:H7 EDL933
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 11214)
Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
Postai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,
Grobeck, E.J., Davis, N.W., Lim, A., Dimallanca, E., Potamousis, K.,
Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,
Welch, R.A. and Blattner, F.R.
Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
NATURE 409 (6819), 529-533 (2001)
21074935
PUBMED 1126551
REFERENCE
AUTHORS Escherichia coli O157:H7 EDL933
Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
Postai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,

```



```

MG1655: B0130"
/codon_start=1
/transl_table=11
/product="orf. hypothetical protein"
/protein_id="AAG54434.1"
/db_xref="GI:12512845"
/transl_table="MYQAVYILLMLFTASVALPARYMOTIENAAVMAOIGDKMWI
VGNIRAGCIITAVEPTAASYAENFGDFKGFIDKGHELPVQCGKQVEDGIDGNKPLSN
QNLVTWKDTPYNAPSVGSAPFGVLADNURYIPILHKLKDLNQTQWIRIGDLRYLS
ALDQKQNGPLQPVUTYHHILRDEENTFRHTSTTSYAFNNQMAWLRDGYATILSMVQ
LEGVQKINLPARAVVITPDDGLKSVSYAYPVLKQYGMKATAFITVSRIKRHPQKW
EPKSLQPMWSVSELNEDVPDFOSHTHELHRVDGVERPILLSRSEHNLFPFASRRA
LAQNPWVHLYSPFGCFNDKAYKAANDAGFHLAVTMTKGKVPKPGNPLLKRLYLRI
TDSLETWSRLVSNVQPGQ"
complement(7611..7991)
/gene="panD"
/note="synonym: Z0142"
complement(7611..7991)
/gene="panD"
/function="enzyme; Biosynthesis of cofactors, carriers:
Pantothenate"
/note="Residues 1 to 126 of 126 are 100.00 pct identical
to residues 1 to 126 of 126 from Escherichia coli K-12
Strain MG1655: B0131"
/codon_start=1
/transl_table=11
/product="aspartate 1-decarboxylase"
/protein_id="AAG54435.1"
/db_xref="GI:12512846"

```

Query Match	9.0%	Score 118.8	DB 1	Length 11214
Best Local Similarity	48.3%	Pred. No. 1.2e-12		
Matches 427	Conservative 0	Mismatches 397	Indels 60	Gaps 1
QY	239	CGCTCGATGGCGCCAGGCGAAGCTCATGGCGCTGCTCCACACATGGCTTCTCCAC	298	
DB	10269	CGCGGCTGGTATGGAAGCAGCGCTGGCGCTGCTACCTACCATGGCAACCTGCAT	10210	
QY	299	GCGGGCCACTCTCGCTCGTGGCCAGGCGCCGCAACTCTCCGACGTCGTGGCGCTCTCC	358	
DB	10209	GATGGCCACATGAAGCTGTGCACGAAGCAAAAGCCGCGCATGTGTGTCGTCACT	10150	
QY	359	ATCTAGCTCAACCGGGCCAAATTCGCCCCACGGAGGACCTCTCACACTACCCCTCGAC	418	
DB	10149	ATTTTCGTTAACCCGATGCAGTTGCAGCCCGCGAGGACCTGCTCGGTACCCACGACC	10090	
QY	419	TTCCAGCGGCGAAGTAAAGAACTTCGGCTCCGTTCCCGGGCGTGCAGCTCGTTTCCAT	478	
DB	10089	TTGCAGGAAGACTGCCAAGCT-----	10067	
QY	479	CCCCGTAACCTTGACGATTACGGGAAGAACGCTGCTGTGTGACGTGGCAGAGCTGGTGA	538	
DB	10066	-----GAACAAGCCAAAGTGGATCTGTTTTCGCGCCCTCG	10030	
QY	539	ATGCTGCTGCGCTTGAGAGTGGTCCGGCGACGAAAGTTGGCTGAGGTTTGAGAGCTG	598	
DB	10029	GTAAGAAGATCTACCCGAACGGTACTGAAACCCATCTTACGTTGACGTTCCAGGGCTT	9970	
QY	599	GAAATGGGCGCTGTGGGAAGACAGGCGCGTTCCTTCAGAGGGGTGGCAGCTGTGGTG	658	
DB	9969	TCGACCATCTTGGAAGGCGCCAGCGCGCGGGGCATTTTCGCGCGCTGTCAGCTATCGTC	9910	
QY	659	ACGAAGTCTGTTAATTTGAGAGCCAGATGGCTGTGTTCCGTGAAGAGGATATCAG	718	
DB	9909	AGCAAGCTGTTAACTGTGTCAGCGCGACATCGCCTGCTTCGTGGTGAAGAGACTTTCAG	9850	
QY	719	CAGTGGCGGCTTATTTCAGAGGATGGTTTCAGATCTTGATTTTCCATAAAGTCATAGGT	778	
DB	9849	CAACTGGCGCTGATCCGCAAAATGTTGCTGATATGGCTTTGATTCAGATTCGCT	9790	
QY	779	GCTGAATAACAGTGATTAATGATGGCTGGCAATGATTCACGTAATGTGCACCTTCA	838	
DB	9789	GTGCCAAATTATCGGCGCAAGACGGTCTGGCACTGAGTTCCTGTAACGGTTATCTGACG	9730	



SLVICVINIGOLPHVWISIFESAFGWQERAGAGYTLQAITNGPORSMPSEACMG  
STENAAAAASWPHPAQGIQVOMIGIFIDTLVICTASAMLIILLAGNTTYMELEGIQ  
LIQKMRVLMSWGAETVILVILFAFSSIVANYIYAENLFFLRNNPKATWLCRIC  
TFATVIGTLLSLPLMWQLADIIMACMAITNLTAILLSPVHTIASDYLRQKLGVR  
VFDPLEYPDIGHQLSRDAWDVDSQE"

Query Match 9.0%; Score 118.8; DB 1; Length 281530;  
Best Local Similarity 48.3%; Pred. No. 1.6e-12;  
Matches 427; Conservative 0; Mismatches 397; Indels 60; Gaps 1;

QY 239 CGTTCGATGGGGCCCAAGCTCATTTGGGTGTCCTCCACCATGGGCTTCTCCAC 298  
DB 153344 CGCGCTGCGTATGGAAGCAAGCGGTGCTGTACTACCATGGGCAACCTGCAT 153285  
QY 299 GCGGGCCACCTCTCGCTGTCGGCCAGCGCCGCCAACTCTCGACGTCGTCGCTCC 358  
DB 153284 GATGGCCACATGAAGCTGTGACGAAGCAAGCCGCGCGATGTGTCGTCACT 153225  
QY 359 ATCTACGTCAACCGGGCCCAATTCCGCCCCAGGAGGACCTCTCCACTACCCCTCGAC 418  
DB 153224 ATTTTCGTTAACCGATGCGAGTCGACCGCCGAGGACCTGGCTGGTACCCACGCC 153165  
QY 419 TTCGACGGCAGCTAAAGAACTCGCTCCGTTCCCGCGCGCTGACGTCGTTTCCAT 478  
DB 153164 TTGAGGAAGACTGCGAAAAGCT----- 153142  
QY 479 CCCGTAACCTTACGATTACGGGAAGAACGTCGTGTCGTCGTCGAGAGGCTGTGGA 538  
DB 153141 -----GAACAGCGCAAAATGGAATCTGGTTTTCGCGCCCTCG 153105  
QY 539 ATGCTGTCGCTGTGAGAGTGGTCCGGCCAGCAAGTTGGTGAGGGTTGAGAAGCTG 598  
DB 153104 GTAAAGAGATCTACCGAAGCGTACTGAACCCATCTTACGTTGACGTTCCAGGCTT 153045  
QY 599 GAATTGGGGCTGTGTGGAAGAGCAGCCCGCTTTCTTACAGGGGTGGGACTGTGGTG 658  
DB 153044 TCACCATGCTGGAAGCGCCAGCGCGCGGGCAATTTCCGCGGTGTCGACTATCGTC 152985  
QY 659 ACGAAGTTGTTAATATTGTGAGGCCAGATGTGGCTGTGTCGTTAAGAGGATTATCAG 718  
DB 152984 AGCAAGCTGTTAACTGTTCCAGCGGACATCGCTGCTTCCGTTGAAAAGGACTTTCAG 152925  
QY 719 CAGTGGCGCTTATTCAGAGGATGTCGAGATCTTGATTTTCCATAAAGTAGT 778  
DB 152924 CAACTGGCTGATCCGCAAAATGGTTGCTGATATGGGCTTTGATATTGAGATTGTCGT 152865  
QY 779 GCTGAATAACAGTGAATATGATGGCTGGCAATGAGTTCACTAATGTGCACCTTTCA 838  
DB 152864 GTGCCAATTATGCGGCCAAAGACGCTCTGGCACTGAGTTCGGTAAACGTTATCTGACG 152805  
QY 839 CCTGAAGAGAGGAAAGGCATCATCAATAAATCATTTGTTAAGAGCTAAATCAGCA 898  
DB 152804 GCAGAAACAGCAAAATTCGCCCGCTCTGTACAAAGTTTAAAGTTTCGATTGCCGACAA 152745  
QY 899 GCAGGAGATGTCAGTGCATTTGAGAGATTGACAAAATTTGGTCTATCCAAAGTTACT 958  
DB 152744 TTGACGCTGGCGAACGGGATCTCGAATGAATTTGCTATTTCGCGGCAAGAACTGAAT 152685  
QY 959 GATGCTGTGGAAGGATGATTTATGCTGAGATTGTTGATCAAAATTAATTTGGAGAAAGTG 1018  
DB 152684 GAAAAGGCTTCGCGCCCGATGATATTCAGATTCGCGATGCCGACACATTTCTGGAGTC 152625  
QY 1019 GAACAGATCAAGATCTCTGCTCTTCTGTTGTCGATGGTTTGGCAAGTCAGGCTT 1078  
DB 152624 TCTGAACACGCAAAACGGGCAAGTAATCTGTTAGCGCTGGCTTGGCGATGCTCGCCTG 152565  
QY 1079 ATAGACACATGGAATCAACTGTCAATGAATGTTGATCTAA 1122  
DB 152564 ATGACACAAAATGTCGAGCTGGGTAATCTTACTACTAA 152521

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 21:18:00 : Search time 3107 Seconds  
(without alignments)  
10294.397 Million cell updates/sec

Title: US-10-033-269-8  
Perfect score: 1316  
Sequence: 1 tcggcagagcgtttctcag.....tggcattgctattgtctagg 1316

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*

1: em\_estha.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estma.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_est3.\*

12: gb\_est4.\*

13: gb\_est5.\*

14: gb\_estfun.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	47.8	629	13	BU965091
2	488.8	37.1	573	13	BQ453747
3	416	31.6	571	12	BM731347
4	346	26.3	1784	11	AY110983

5	309.8	23.5	541	13	BU034064
6	303.2	23.0	676	10	EG098937
7	288.2	21.9	576	13	BU030691
8	283.8	21.6	625	13	BU031672
9	283	21.5	503	13	BQ976077
10	276.6	21.0	531	9	AI730340
11	259.2	19.7	616	14	CB629807
12	254.4	19.3	545	9	AU290923
13	251.4	19.1	746	14	CB629808
14	247	18.8	875	29	BZ547854
15	244	18.5	838	10	BF064760
16	242.2	18.4	651	14	CB879036
17	241.4	18.3	573	9	AU286177
18	230.4	17.5	625	13	BQ768349
19	219	16.6	610	28	BH505138
20	218.2	16.6	575	29	BZ408845
21	203.8	15.5	425	9	AW719829
22	191.8	14.6	577	14	CB091543
23	190	14.4	488	9	AL829512
24	162.6	12.4	455	9	AU286178
25	160	12.2	606	9	AI998016
26	153.4	11.7	450	14	CA687321
27	144	10.9	404	14	CA725201
28	141.4	10.7	215	10	BF153835
29	138.8	10.5	779	29	BZ408852
30	127	9.7	225	13	BQ765881
31	124.2	9.4	868	10	BF267302
32	118	9.0	571	13	BU049399
33	114.8	8.7	545	12	BM276865
34	110.2	8.4	516	12	BM419808
35	106.6	8.1	518	12	BM276771
36	105.6	8.0	495	29	AG222917
37	104.4	7.9	439	12	BI203409
38	101.6	7.7	563	14	CA758922
39	100	7.6	564	10	BE040497
40	96.4	7.3	1065	29	CNS078AS
41	95	7.2	725	29	BZ560200
42	91.8	7.0	988	29	BZ573502
43	88.6	6.7	477	12	BM259201
44	83.4	6.3	581	13	BQ590858
45	83	6.3	451	9	AU173408

ALIGNMENTS

RESULT 1  
BU965091  
LOCUS  
DEFINITION  
BU965091 629 bp mRNA linear EST 21-OCT-2002  
sat06b11.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl036-12765 5' similar to TR:024035 024035  
PANTOATE--BETA-ALANINE LIGASE ;, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BU965091.1 GI:24205838  
EST.  
Glycine max (soybean)  
Glycine max  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE  
AUTHORS  
1 (bases 1 to 629)  
Shoenaker R., Keim P., Vodkin L., Erpelding J., Corryell V., Khanna  
A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C.,  
Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers  
Y., Person B., Swaller F., Gibbons M., Pape D., Harvey N., Schurk  
R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann  
R., Waterston R. and Wilson R.  
Public Soybean EST Project  
Unpublished  
JOURNAL  
COMMENT  
Contact: Shoenaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine

BU034064 QHJ2P22.Y  
EG098937 EST463456  
BU030691 QHJ16C05  
BU031672 QHJ19C05  
BQ976077 QHJ20G07  
AI730340 BNGH41668  
CB629807 OSIIB06F  
AU290923 AU290923  
CB629808 OSIIB06F  
BZ547854 OGAH462TC  
BF064760 HV\_CBB001  
CB879036 HFI0117T  
AU286177 AU286177  
BQ768349 EBR008 SQ  
BH505138 BOHNP38TR  
BZ408845 OGAAM35TC  
AW719829 LJNBS10F  
CB091543 he93f02.g  
AL829512 AL829512  
AI998016 701672578  
CA687321 wlm96.pk0  
CA725201 wds3f.pk0  
BF153835 030F05 Ma  
BZ408852 OGAAM35TM  
BQ765881 EBR003 SQ  
BF267302 HV\_CBA001  
BU049399 1111007E0  
BM276865 952009C02  
BM419808 RO19A11 O  
BM276771 952010G01  
AG222917 Lotus Jap  
BI203409 EST521449  
CA758922 OS08F10-T  
BE040497 BE07F10 O  
AL433754 T3 end of  
BZ560200 pacs2-164  
BZ573502 msh2 3228  
BM259201 952010G01  
BQ590858 E012596-0  
AU173408 AU173408



4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccu@resgen.com web site:  
 www.resgen.com

Seq primer: -40RP from Gibco  
 High quality sequence stop: 446.

#### FEATURES

source

1. 629

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-12765"

/tissue\_type="somatic embryos cultured on MSD 20"

/lab\_host="DH10B"

/clone\_lib="Gm-cl036"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This  
 cDNA library was constructed from mRNA isolated from  
 somatic embryos (age ranging from 2 months to 9 months)  
 cultured on MSD 20. The library was prepared using the  
 Life Technologies pSuperScript cDNA library construction  
 kit. Complementary DNA was synthesized from mRNA using a  
 poly (dT) sequence with a NotI restriction site. SalI  
 linkers adapters were ligated to the blunt-ended cDNA  
 fragments followed by NotI digestion. The cDNA fragments  
 were directionally cloned into the NotI-SalI restriction  
 site of the pSPORT1 vector. The ligated cDNA fragments  
 were transformed into E.coli ElectroMax DH10B host cells.  
 This library was constructed in the laboratory of Dr. Lila  
 Vodkina by Anu Khanna at the University of Illinois at  
 Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

BASE COUNT

124 a 197 c 173 g 135 t

#### Query Match

Best Local Similarity 47.8%; Score 629; DB 13; Length 629;

Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 39 ACACACCAAGAGTACTTCAACCCAGAGTCATATACTTGCCTAACACCTACCTCT 98
Db 1 ACACACCAAGAGTACTTCAACCCAGAGTCATATACTTGCCTAACACCTACCTCT 60

QY 99 TCAACACCGGTTTCGGCCATCAGCTTCGTCCTTCTTCTTATTAATCTCTCAGAAAC 158
Db 61 TCAACACCGGTTTCGGCCATCAGCTTCGTCCTTCTTCTTATTAATCTCTCAGAAAC 120

QY 159 TCCCACTGTTTCCAAATATTATTCAGATGGTTCAGCCCCCAGGGTGATCTCCGACAAAG 218
Db 121 TCCCACTGTTTCCAAATATTATTCAGATGGTTCAGCCCCCAGGGTGATCTCCGACAAAG 180

QY 219 CTTGATCGGAGCTGTCGGCTGATGCGGGCCAGGGCCAGGTCATTTGGTGTGTC 278
Db 181 CTTGATCGGAGCTGTCGGCTGATGCGGGCCAGGGCCAGGTCATTTGGTGTGTC 240

QY 279 CCACATGGGTTCTTCCACGCGGGCCACCTCTGCTGTCGCCCCAGGCCGCCCACTCT 338
Db 241 CCACATGGGTTCTTCCACGCGGGCCACCTCTGCTGTCGCCCCAGGCCGCCCACTCT 300

QY 339 CCGAGCTGTCGCGCTTCCATCTACGTCAACCCCGGGCCAAATTCGCCCCCGAGGAGCC 398
Db 301 CCGAGCTGTCGCGCTTCCATCTACGTCAACCCCGGGCCAAATTCGCCCCCGAGGAGCC 360

QY 399 TCTCCACTACCCCTCCGACTTCGACCGGCGAGCTAAGAACTCGCTTCGTCGCGG 458
Db 361 TCTCCACTACCCCTCCGACTTCGACCGGCGAGCTAAGAACTCGCTTCGTCGCGG 420

QY 459 CGCTCGAGCTGCTTTTCCATCCCGTAACTTGTACGTTACGGGAAGAACGGTGTGTG 518
Db 421 CGCTCGAGCTGCTTTTCCATCCCGTAACTTGTACGTTACGGGAAGAACGGTGTGTG 480

```

```

QY 519 ACGTGGCAGAGCTGGTGAATGGTGTGCTGCTTGAAGTGGTGGTCCGGCAGCAAGTT 578
Db 481 ACGTGGCAGAGCTGGTGAATGGTGTGCTGCTTGAAGTGGTGGTCCGGCAGCAAGTT 540

QY 579 GGTGAGGTTGAGAGCTGGAATGGGCTGTGTGGAAAGAGCAGCCCGTTTCTTCA 638
Db 541 GGTGAGGTTGAGAGCTGGAATGGGCTGTGTGGAAAGAGCAGCCCGTTTCTTCA 600

QY 639 GAGGGTGGCGACTGTGTGACGAAGTTG 667
Db 601 GAGGGTGGCGACTGTGTGACGAAGTTG 629

```

#### RESULT 2

BQ453747/c

LOCUS

DEFINITION

Gm-cl081-2478 5' similar to TR:024035 024035 PANTOATE--BETA-ALANINE

LIGASE ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine

1 (bases 1 to 573)

Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Corvelli,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco

High quality sequence stop: 474.

Location/Qualifiers

1. 573

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl081-2478"

/tissue\_type="Roots of 7 day old 'Bragg' seedlings"

/dev\_stage="7 days old"

/lab\_host="DH10B"

/clone\_lib="Gm-cl081"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site\_2:

XhoI; The mRNA was isolated from roots of 7 day old

'Bragg' seedlings that were mock-infected 48 hours prior

to harvest. Dr. Gary Stacey generously donated the

tissue. The roots were flash-frozen in liquid nitrogen.

Stratagene's cDNA Synthesis kit (catalog number 200401)

was used to synthesize the cDNA. First-strand synthesis

was performed with 5-methyl dCTP, hence the ligated cDNA

was hemimethylated. A modification of Stratagene's

first-strand synthesis primer was used. An 'anchor'

nucleotide (V=A, C, or G) was added to the 3' end of the

primer [GAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor

the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University." 93 g 185 t 1 others

```

BASE COUNT      175 a 119 c      93 g 185 t 1 others
ORIGIN
Query Match      37.1%; Score 488.8; DB 13; Length 573;
Best Local Similarity 99.4%; Pred. No. 7.7e-126;
Matches 490; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 824 AATGTGCACCTTCACTGAAGAGAGGAAAGGCACTATCAATAAATAATCATTTGTTA 883
Db 573 AATGTGCACCTTCACTGAAGAGAGGAAAGGCACTATCAATAAATAATCATTTGTTA 514
QY 884 AGAGCTAAATCAGACAGAGAGTGTGAGTGCATTTGAGAAAGTTGCAAAATTTGGTC 943
Db 513 AGAGCTAAATCAGACAGAGAGTGTGAGTGCATTTGAGAAAGTTGCAAAATTTGGTC 454
QY 944 ATCCAAAGTGTTACTGATGCTGTGGAAGGATGATGCTGAGATTTGTCGATCAAAAT 1003
Db 453 ATCCAAAGTGTTACTGATGCTGTGGAAGGATGATGCTGAGATTTGTCGATCAAAAT 394
QY 1004 AATTGCGAAAGTGAACAGATCAAGAGTCCCTGCTCTTCTGTGCTGCGATGGTTT 1063
Db 393 AATTGCGAAAGTGAACAGATCAAGAGTCCCTGCTCTTCTGTGCTGCGATGGTTT 334
QY 1064 GCGAAGTCAGGCTTATAGACACATGGAATCAACTTGTCAATGAATTTGATCTAAC 1123
Db 333 GCGAAGTCAGGCTTATAGACACATGGAATCAACTTGTCAATGAATTTGATCTAAC 274
QY 1124 CTTCGTGATCTCAACATGCGGCACATGCTTAATTAATAGTTGCGGCCACGCTTAAC 1183
Db 273 CTTCGTGATCTCAACATGCGGCACATGCTTAATTAATAGTTGCGGCCACGCTTAAC 214
QY 1184 AATTCACAGTTCATGGTTATAGTCATGACAAATTTTTTTTCTGCCAGCCATACATGAT 1243
Db 213 AATTCACAGTTCATGGTTATAGTCATGACAAATTTTTTTTCTGCCAGCCATACATGAT 154
QY 1244 TACTTTAGATGCAATTTACCGATCATAAATTTCTATGAGAGCTGCTACCGATGCGATT 1303
Db 153 TACTTTAGATGCAATTTACCGATCATAAATTTCTATGAGAGCTGCTACCGATGCGATT 94
QY 1304 GCTATTGCTAGG 1316
Db 93 GCTATTGCTAGG 81

RESULT 3
BM731347
LOCUS
DEFINITION
sal70909.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl061-4553 5' similar to TR:024035 024035 PANTOATE--BETA-ALANINE
LIGASE i, mRNA sequence.
ACCESSION
BM731347
VERSION
BM731347.1 GI:19052680
KEYWORDS
EST.
SOURCE
Glycine max (soybean)

```

## ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurossids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine

## REFERENCE

## AUTHORS

1 (bases 1 to 571)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

## TITLE

## JOURNAL

## COMMENT

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 426.

Location/Qualifiers

## FEATURES

source

1..571

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl061-4553"

/tissue\_type="mature flowers of field grown plants"

/lab\_host="DH10B"

/clone\_lib="Gm-cl061"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from mature flowers of field grown plants for the cultivar

Raiden. Complementary DNA was synthesized from mRNA using

a primer consisting of a poly(dT) sequence with a XhoI

restriction site. EcoRI adapters were ligated to the

blunt-ended cDNA fragments followed by XhoI digestion. The

cDNA fragments were directionally cloned into the

EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host

cells (GibcoBRL). This library was constructed in the

laboratory of Dr. Randy Shoemaker."

167 a 93 c 126 g 185 t

## BASE COUNT

## ORIGIN

Query Match

31.6%; Score 416; DB 12; Length 571;

Best Local Similarity 100.0%; Pred. No. 2.2e-105;

Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

```

Db      241  ATGGGCCACATGCTTAATTAATAGTTCGGGCCACAGTGTAAACAATTCTAACAGTTCATG 300
QY      1201  GTTATAGTCATGACAAATTTTTTTTCTGCGACCCATACATGATTACTTTGTAGATGCAITTT 1260
Db      301  GTTATAGTCATGACAAATTTTTTTTCTGCGACCCATACATGATTACTTTGTAGATGCAITTT 360
QY      1261  TACCCGATCATAAATTTCTATGAGAGCTGCTACACAGTGCATGCTATTTGTGAGG 1316
Db      361  TACCCGATCATAAATTTCTATGAGAGCTGCTACACAGTGCATGCTATTTGTGAGG 416

RESULT 4
LOCUS   AY110983                1784 bp      mRNA      linear      HTC 17-OCT-2002
DEFINITION   Zea mays CL778.1 mRNA sequence.
ACCESSION   AY110983
VERSION     AY110983.1  GI:21215573
KEYWORDS    HTC.
SOURCE      Zea mays
ORGANISM   Zea mays
REFERENCE   1  (bases 1 to 1784)
AUTHORS    Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
JOURNAL    Unpublished (2002)
REFERENCE   2  (bases 1 to 1784)
AUTHORS    Coe,E.H.
TITLE      Direct Submission
JOURNAL    Submitted (25-Apr-2002) Maize Mapping Project, University of
            Misouri, Columbia, MO 65211, USA
COMMENT    If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.

FEATURES             Location/Qualifiers
     source           1..1784
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:632883"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"

BASE COUNT  367 a 393 c 405 g 390 t 229 others
ORIGIN

Query Match      26.3%; Score 346; DB 11; Length 1784;
Best Local Similarity 57.8%; Pred. No. 1.9e-85;
Matches 534; Conservative 0; Mismatches 375; Indels 15; Gaps 2;

QY      190  TCCAGCCCCAGGTGTATCTCCGACAGGCTCGATCGGAGCTGGTCGCGCTCATGCG 249
Db      250  TCATGAGCCGAGGTGTATCTCCGACAGGCGCGANNNNNNNNNNNNNNNNNNNNNNNNN 309
QY      250  G3CCCAGGCGAGCTCATTGGCTGGTCCGCCACCATGGCTTCCTCCAGCGGGCCACCT 309
Db      310  NNNNNNNNNNGGCGCGCTTTAGTCCCCACGATGGGCTTCCTCCAGGAGGACACCT 369
QY      310  CTCGCTCGT-----GGCCAGCGCCGCAACTCTCCGACGTCGTCGCGCTCCAT 360
Db      370  CTCGCTCGTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 429

```

```

QY      361  CTAGCTCAACCCGGG3CCAAATTCGCCCCCAACGAGGACCTCTCCACCTACCCTCCGACTT 420
Db      430  CTAGCTCAACCCGACGAGTTCGNNNNNACGAGGACCTCGCCACCTACCCTCCGACTT 489
QY      421  CGAGCGGACGCTAAAGAAATTCGCGTCCGTTCCCGGGGCGTGACGTCGTTTCCATCC 480
Db      490  CGCGGGGACCTCGCAAGCTGGCGCGCACCGGGGTGTCGCCGCAAGTATTTTGTNNNN 549
QY      481  CCGTAACTTGTACGATTACGGGAAGACGGTGGTGGTGACGCTGGCAGAGGCTGGTGAAT 540
Db      550  GGACCTCTTACGTCGCGGACGCGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 609
QY      541  GGTGTCTGTCGCTTGAG-----AGTGGGTCGGGCACGAAAGTTGGGTGAGGGTGGAGAA 594
Db      610  NNNNNNTTGCCTGAGGATGAGGTGGGCATGCGCACGAGACGTGGATTCGGGTGGAGCG 669
QY      595  GCTGGAATGGGGCTGTGTGGGAAGACAGCCCGTTTCTTCAGAGGGGTGGCGACTGT 654
Db      670  GCTGGAGAAGGGGCTGTGCGGGAGCAGCAGGCCCTGTCTTCTTCGCGGCGTGGCCACCGT 729
QY      655  GGTGACGAAGTTGTTAATATTGTGGAGCCAGATGTGGCTGTGTTCCGTAAGAAGAGTATA 714
Db      730  GGTGCGCAAGCTGTTCATATATCGTGGAGCGGACGTCGCCGTGTTCGGGAGAGAGTATA 789
QY      715  TCAGCAGTGGCGGCTTATTTCAGAGGATGGTTTCGAGATCTTTCGATTTTCCATAAAGTCAT 774
Db      790  TCACGAGTGGCGGCTCAITTCGAGTGGTTCGATCTCGATTTTGCATACAGATAGT 849
QY      775  AGTGTGTAATAACACGTGATATGATGGCTCGGCAATGAGTTCAGTATGTGCACCT 834
Db      850  TGGTTCGAGGTAGTGGCAGAAAGCTGATGCTTTCGCAATGAGCTCTCGCAATGTAATCT 909
QY      835  TTCACCTGAAGAGAGGAAAGGACACTATCAATAAATAATCAATTCATGTAAGAGCTAAATC 894
Db      910  GTCAGAGGAGATAGAAAGAGGGCTTATCGATCAGTAGATCGCTGTTGGTATGCTAGAAC 969
QY      895  AGCAGCAGGAGATGGTCAGGTGCTATTGTGAGAGCTTCACAAATTTGTCATCCAAAGTGT 954
Db      970  CGCGCCCTCAGTGGAGCAACCGTACCCAGAGATAAAGATCAATATGTCGCGACAAT 1029
QY      955  TACTGATGCTGGTGGGAAGATTCGATATGCTGAGATTTGATCAAAATATTTGGAGAA 1014
Db      1030  TACAGAAGCTGGCGGTGAGGTGACTATGATTTGAGATTGGGGCAGAAAGCTTGGTGCC 1089
QY      1015  AGTGGACAGATCAAGAGTCTCTGTCGCTTCTGTGTTGCTGCTGATGTTGGCAAGTCAG 1074
Db      1090  TGTGGAGAGGATGACCGCCCTTGTGTAATTTGTGTCGCGCATGGTTTGGAAAGGTGAG 1149
QY      1075  GCTTATAGACAAACATGGAATCAA 1098
Db      1150  GCTAATTGACAAACATCGAGATCCA 1173

RESULT 5
LOCUS   BU034064                541 bp      mRNA      linear      EST 23-AUG-2002
DEFINITION   QJ2P22.VG.ab1 QH_BFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QJ2P22, mRNA sequence.
ACCESSION   BU034064
VERSION     BU034064.1  GI:22469584
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
ORGANISM   Helianthus annuus
REFERENCE   1  (bases 1 to 541)
AUTHORS    Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J., Ellison
            P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.

```

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compgenomics.ucdavis.edu/  
 JOURNAL Unpublished  
 COMMENT Contact: Alexander Kozik [R.W.Michelmores]  
 Department of Vegetable Crops, R.W.Michelmores Lab  
 University of California at Davis (UCD)  
 Amundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michelmores@vegmail.ucdavis.edu]  
 belongs to contig QH\_CA\_Contig3643, see http://cgpdb.ucdavis.edu/  
 for details.  
 Plate: QH22 row: P column: 22.

FEATURES Location/Qualifiers  
 source  
 1..541  
 /organism="Helianthus annuus"  
 /mol\_type="mRNA"  
 /cultivar="RHA280"  
 /db\_xref="taxon:4232"  
 /clone="QH22P22"  
 /lab\_host="E.coli"  
 /clone\_lib="QH\_EFGHJ sunflower RHA280"  
 /note="Vector: pBRCN4S1A8; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG\_LIB=QH\_EFGHJ sunflower RHA280  
 TAG\_SEQ=germinating seeds  
 TAG\_SEQ=TCTGTGCGGG  
 BASE COUNT 152 a 171 c 176 g 142 t  
 ORIGIN

Query Match 23.5%; Score 309.8; DB 13; Length 541;  
 Best Local Similarity 73.6%; Pred. No. 1.5e-75;  
 Matches 395; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
 QY 568 GCACGAAGTGGGTGAGGCTTGAGAGCTGGAATGGGCTGTGGGAGAGAGCGCC 627  
 Db 1 GCATGACATGGGTAGAGTTGAGAGCTGGAGAAAGGATGTGTGGGAGAGTAGGCC 60  
 QY 628 CGTTTCTTCAGAGGGGTGGCGACTGTGTGACGAAGTGTGTTAATTTCTGAGCCAGA 687  
 Db 61 TGTGTTTTTTAGAGGGGTGGCTACTGTGTGGCTAGTGTGTTAATTTGTAACCGGA 120  
 QY 688 TGTGGCTGTGTCGGTAAAGAGGATTATCAGAGTGGCGCTTATTTCAGAGGATGTTGC 747  
 Db 121 TGTTCGGGTTTTTGGGAAGAGGATTATCAGCAGTGGCGGTTATTTCAGCGGATGTTGC 180  
 QY 748 AGATCTTGATTTTCCATAAAGTGTAGTGTGCTGAATGAATACACGATGATTAATGTCCT 807  
 Db 181 AGATCTTGATTTTCTGTTTAAAGTGTGTTGTTCTGAGTTGAACGAGAGAGATGCGCT 240  
 QY 808 GGCATGAGTTACGTAATGTGACACCTTTTCACCTGAAGAGAGGAGAAAGSCATATCAAT 867  
 Db 241 GGCATGAGTCCGCAACGTTGATCTTTTCACACAGAGAGAGGAAACAGGCCCTTTCGAT 300  
 QY 868 AATAAATCATTTTAAAGCTAAATCAGCAGCAGGAGATGTTTCAGTGCATTGTGAGAA 927  
 Db 301 AAGCCGGGCTTTATTGAAGCCAAATCAGCTGCCAAAAGGGCAATAGCAATTCAGTAA 360  
 QY 928 GTTGACAAATTTGGTTCATCAAGTGTACTGATGCTGGTGAAGATCAATATGCTGA 987  
 Db 361 GTTGAGAGATCATGTGATCAAAAGTACAGAAAGCAGCGGCTGGAAAAATAGATTATGCTGA 420  
 QY 988 GATTGTTGATCAAAATAATTTGAGAAAGTGGAAACAGATCAAGAGTCCTGTGCTCTTCG 1047  
 Db 421 GATTGTGATCAAGAAGTCTAGAAACGGTTGAAGAGATCAGGATGGAGTTGTGGTGTG 480

QY 1048 TGTGTGTCATGGTTTGGCAAGTCAAGCTTATAGACAAATGAAATCACTTGTG 1104  
 Db 481 TGTGTGTCATGGTTTGGCAAGTCAAGCTTATAGACAAATGAAATGATGCTG 537  
 RESULT 6  
 BG098937/c 676 bp mRNA linear EST 07-MAR-2003  
 LOCUS EST463456 sprouting eyes/shoots Solanum tuberosum cDNA clone  
 DEFINITION CSTC8020 3' sequence, mRNA sequence.  
 ACCESSION BG098937  
 VERSION BG098937.1 GI:12588972  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 676)  
 AUTHORS van der Hoeven,R.S., Bezzerides,J., Cho,J., Utterback,T., Hansen  
 ,C.L., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker  
 ,B.  
 TITLE Generation of ESTs from potato sprouting eyes/shoots  
 JOURNAL Unpublished  
 COMMENT Contact: Robin Buell  
 The Institute for Genomic Research  
 712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/.

FEATURES Location/Qualifiers  
 source  
 1..676  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="CSTC8020"  
 /tissue\_type="sprouting tubers"  
 /dev\_stage="12-14 weeks post harvest"  
 /lab\_host="SOLR"  
 /clone\_lib="sprouting eyes/shoots"  
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Various sizes of sprouting eyes (2 mm to 15 mm) were  
 taken from tubers. The tubers were incubated at 26C in  
 the dark for 2-3 weeks prior to sprouting. The eyes were  
 frozen in liquid nitrogen immediately upon removal from  
 tubers."  
 BASE COUNT 205 a 173 c 92 g 206 t  
 ORIGIN  
 Query Match 23.0%; Score 303.2; DB 10; Length 676;  
 Best Local Similarity 75.8%; Pred. No. 1.2e-73;  
 Matches 388; Conservative 0; Mismatches 123; Indels 1; Gaps 1;  
 QY 597 TGAATTTGGGCTGTGTGGAGAGCAGCCGCTTTCTTCAAGGGGTGGCAGTGTG 656  
 Db 675 TGGAAAGGATGTGTGGGAATAGTAGCCCTGTTTCTTAAAGGGTGTCTACTGTG 616  
 QY 657 TGACGAAGTTGTTTAATATTGTGAGCCAGATGTGCTGTGTTCCGTAAGAAGGATTATC 716  
 Db 615 TTACCAAGTTGTTCAATATAGTTGAGCCTGATGTTGTCTTTGGTAAGAAGGATTATC 556  
 QY 717 ACCAGTGGCGCTTATTTCAGAGATGGTTCGAGATCTTGAATTTTCCATAAAGTGATAG 776  
 Db 555 AGCAATGGAGGATTATACAGAAATCGTCAGAGATCTTGATTTTGGGATAAAGGTGATTG 496  
 QY 777 GTGCTGAATAACAGCTGATGATGCTGGCATGAGTTTCAGTTATGTGCACCTTT 836  
 Db 495 GTTCTGAATATGACAGAGCATGATGCGCTTCGAATGAGTTCCCGCTAATGTGAACCTTT 436  
 QY 837 CACCTGAAGAGAGGGAAGGCACTATCAATAAATAAATCATTTGTTTAAGAGCTAATCAG 896

```

Db      435 CACCTCGGACAGACAAAGGGCTTATCAATTAGTCGTGGCTGTCTAGACACAAGTTG 376
QY      897 CAGCAGGAGATGTCAGGTGCATGTGAGAGTTGACAAATTTGGTCAATCCAAAGTGTTA 956
Db      375 AAGCAGAAAGGGTCAGGTCAACTGCGAGGAGCTGATAAATACCTGCCATCTTACCATAT 316
QY      957 CTGATCTCGTGAAGGATCGATTATGCTGAGATGTTGATCAAAATAATTTGGAGAAAG 1016
Db      315 CTGAAGCTGTTGAACGGTTGATTATGCTGAGATTGTGATCAAGAAAGTTTAGAGCCAG 256
QY      1017 TGAACAGATCAAGAGTCTGCTGCTTCTGCTGTTG-CTGCAATGTTTGGCAACTCAGG 1075
Db      255 TGAACAACATCAAGAGACCAAGTTGTTTGTAGCCAGTAGTGGTTTGGAAAGGTGAG 196
QY      1076 CTTATAGACAACATGGAATCAACTTGTCAAT 1107
Db      195 CTTATTGACAACATGGAATAGATGATATAAT 164

RESULT 7
BU030691
LOCUS
DEFINITION
  QHJ16C05.Yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
ACCESSION
  BU030691
VERSION
  BU030691.1 GI:22466211
KEYWORDS
  EST.
SOURCE
  Helianthus annuus (common sunflower)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    asterids; campanulids; Asterales; Asteraceae; Asteroideae;
    Heliantheae; Helianthus.
  1 (bases 1 to 576)
  Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
  Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
  ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
  Church,S., Jackson,L. and Bradford,K.
  Lettuce and Sunflower ESTs from the Compositae Genome Project
  http://compgenomics.ucdavis.edu/
  Unpublished
  Contact: Alexander Kozik [R.W.Michelmore]
  Department of Vegetable Crops, R.W.Michelmore Lab
  University of California at Davis (UCD)
  Asmundson Hall, UCD, Davis, CA 95616, USA
  Tel: 1-(530)-742-1742
  Fax: 1-(530)-752-9659
  Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
  belongs to contig QH_CA_Contig3643, see http://cgpdb.ucdavis.edu/
  for details.
  Plate: QHJ16 row: C column: 05.
  Location/Qualifiers
    1..576
    /organism="Helianthus annuus"
    /mol_type="mRNA"
    /cultiivar="RHA280"
    /db_xref="taxon:4232"
    /clone="QHJ16C05"
    /lab_host="E.coli"
    /clone_lib="QH_EFGHJ sunflower RHA280"
    /note="Vector: pBRCDNASFIAB; The library was constructed
    from 11 different sources of RNA from a single genotype.
    Separate cDNAs were generated using primers that
    incorporated unique 5' and 3' tags to distinguish each
    source of RNA. cDNAs were then pooled, size-fractionated,
    directionally cloned into a custom medium-copy vector and
    transformations made with four size classes to minimize
    size bias. Details of each source of RNA and library
    construction can be obtained at http://cgpdb.ucdavis.edu/
    TAG_L1B-QH_EFGHJ sunflower RHA280
    TAG_TISSUE=germinating seeds
    TAG_SEQ=TCTGTGCGGG"
  163 a 76 c 183 g 154 t

```

## ORIGIN

```

Query Match      21.9%; Score 288.2; DB 13; Length 576;
Best Local Similarity 72.8%; Pred. No. 1.8e-69;
Matches 385; Conservative 0; Mismatches 143; Indels 1; Gaps 1;

QY      568 GCACGAAAGTTGGGTGAGGGTTGAGAACTGGAATTTGGGGCTGTGTGG- GAAGAGACAGC 626
Db      1 GCATGAGACATGGGTAGAGTTGAGAGCTGGAGAAACGAGATGTGTGAGAGTAGGC 60
QY      627 CCGTTTCTTCAGAGGGGTGGCGACTGTGTGACGAAAGTTGTTTAATAATTTGGAGGCAG 686
Db      61 CTGTGTTTTTAGAGGGGTGGCTACTGTGTGGCTAAGTTCTTTAATAATTTGAAACCG 120
QY      687 ATGTGGCTGTCTCGTTAAGAGGATATACAGCAGTGGCGCTTATTCAGAGGATGGTTC 746
Db      121 ATGTTCGGTTTTTGGGAAGAGGATATACAGCAGTGGCGGTTATTTCGACGGATGGTTC 180
QY      747 GAGATCTTGAATTTTCCATAAAAGTATAGTGTCTGMAATAAACACGTGATATATGATGCC 806
Db      181 GAGATCTTGAATTTTCTGTGTTAAAGTGATTTGTTCTGAGTTCAGCGAGAGAAACATGCC 240
QY      807 TGGCAATGAGTTCACGTAATGTGCACCTTTTCACCTGAAGAGAGGAAAGGCACATCAAA 866
Db      241 TGGCAATGAGTCCCGCAACCTGTCATCTTTTCAACAAGAGAGAGGAAACAGGCTTGTGGA 300
QY      867 TAAATAAATCATTTGTTAAGAGCTAAATCAGCAGCAGGAGATGTTTCAGTGCATTTGTGAGA 926
Db      301 TAAGCCGGCTTTTATTTGAGCCAAATCAGTCGCAAAAAGGGCAATAGCAATTTGCAGTA 360
QY      927 AGTTGACAAATTTGGTTCATCCAAAGTGTCTACTGTGTGTGGAGAGATCCATTATGCTG 986
Db      361 AGTTGAGAGATCAGTGTATCGAAAGTATAGAGCAACCCGGTGGAAAATAAGATTATGTTG 420
QY      987 AGATTGTCATCAAAATAATTTGGAGAAAGTGGAAACAGATCAAGAGTCTCTGCTCTCT 1046
Db      421 AGATTGTCATCAAAAGAGGCTAGAAACGGTTGAAGAGATCAAGATGAGTGTGGNGT 480
QY      1047 GTGTGTGTCATGTTTGGCAACTCAGGCTTATAGCAACATGGAAT 1095
Db      481 GTGTGTCAGCATGTTTGGGAAGTGCAGTTGATCGACACATTTGGAAT 529

```

```

RESULT 8
BU031672
LOCUS
DEFINITION
  QHJ19C05.Yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
ACCESSION
  BU031672
VERSION
  BU031672.1 GI:22467192
KEYWORDS
  EST.
SOURCE
  Helianthus annuus (common sunflower)
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    asterids; campanulids; Asterales; Asteraceae; Asteroideae;
    Heliantheae; Helianthus.
  1 (bases 1 to 625)
  Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
  Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
  ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
  Church,S., Jackson,L. and Bradford,K.
  Lettuce and Sunflower ESTs from the Compositae Genome Project
  http://compgenomics.ucdavis.edu/
  Unpublished
  Contact: Alexander Kozik [R.W.Michelmore]
  Department of Vegetable Crops, R.W.Michelmore Lab
  University of California at Davis (UCD)
  Asmundson Hall, UCD, Davis, CA 95616, USA
  Tel: 1-(530)-742-1742
  Fax: 1-(530)-752-9659
  Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
  belongs to contig QH_CA_Contig3643, see http://cgpdb.ucdavis.edu/

```

## FEATURES

## source

for details.

Plate: QH19 row: C column: 05.

FEATURES  
source

Location/Qualifiers  
1. .625  
/organism="Helianthus annuus"  
/mol\_type="mRNA"  
/cultivar="RHA280"  
/db\_xref="taxon:4232"  
/clone="QH19C05"  
/lab\_host="E.coli"  
/clone\_lib="QH\_EFGHJ sunflower RHA280"  
/note="Vector: pBRCNAsf1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/  
TAG\_LIB=QH\_EFGHJ sunflower RHA280  
TAG\_TISSUE=germinating seeds  
TAG\_SEQ=TCTGTGCGG"

BASE COUNT 179 a 188 g 179 t 3 others

ORIGIN

Query Match 21.6%; Score 283.8; DB 13; Length 625;  
Best Local Similarity 72.6%; Pred. No. 3.2e-68;  
Matches 392; Conservative 0; Mismatches 145; Indels 3; Gaps 2;

QY 568 GCACGAAGTTGGGTGAGGTTGAGAAAGCTGGAAATGGGGCTGCTGTTGGCTAGTGTGTTAATTTGTGAGCC 626  
Db 1 GCATGACATGGGTTAGATTGAGAGCTGGAGAAAGGATGTTGGGGAAGAGTANGC 60  
QY 627 CC--GTTTTCTTCAGAGGGTGGCGACTGGTGGTGCACGAAGTTGTTAATTTGTGAGCC 684  
Db 61 CTTGTGTTTTTAGAGGGTGGCTACTGTGGCTAGTGTGTTAATTTGTGAGCC 120  
QY 685 AGATGGCTGTGTTTCGTTAAGAGGATATACAGCTGGCGCTTATCAGAGGATGTT 744  
Db 121 GGATGTTGCGGTTTTTGGGAAGAAGGATATACAGCTGGCGGTTATTCGACGGATGTT 180  
QY 745 TCAGATCTTGATTTTCCATAAAGTGATAGGTGCTGAATACACGCTGATATGATGG 804  
Db 181 TCAGATCTTGATTTTTCGTTAGTGTGATGTTGTTCTGAGTTGAAGCGAAGAAGATGG 240  
QY 805 CCTGGCAATGAGTTTCACTAATGTGCACCTTTTCACTTGAAGAGAGGAAAGGCATATC 864  
Db 241 CCTGGCAATGAGTCCCGCACTGCTATCTTTCACACAGAGAGGAAAGGCATATC 300  
QY 865 AATAAATAATCTTTTAAGAGCTAATCAGCAGCAGGAGATGGTCAGGTGCATTTGTA 924  
Db 301 GATAAGCCGGGCTTTTATTTGAAGCCANATCAGCTGCCAAAAGGGCAATAGCAATTGCAG 360  
QY 925 GAAGTTGACAAATTTTGTGTCATCCAAAGTGTTACTGATGCTGTTGGGAGGATCATATGC 984  
Db 361 TAAATTGAGAAGATTCAGTGATCGAAAGTATAGAAGCGCGGTGGAAAAATGATATGTT 420  
QY 985 TGAGATTTTGATCAAAATAATTTGGAGAAGTTGGAACATCAAGAGTCTCTGCTGCTT 1044  
Db 421 TGAGATTTGATCAGAAGCTTGAACCGTTGAAGATCAAGATGGAGTTGTGGT 480  
QY 1045 CTGTGTTGTCAGGTTTGGCAAGTCAGGCTTATAGACAAATGACATCAATCAACTTGTG 1104  
Db 481 GTGTGTTGACGATGTTTGGGAAGGTGAGGTTGATCGGATGATCGGATGATCGGCTC 540

RESULT 9

BQ976077

LOCUS

DEFINITION BQ976077, yg.ab1 QH ABCDI sunflower RHA801 Helianthus annuus cDNA clone QH120G07, mRNA sequence. EST 21-AUG-2002

ACCESSION

BQ976077

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BQ976077.1 GI:22393600  
EST.  
Helianthus annuus (common sunflower)  
Helianthus annuus

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 503)  
Kozik, A., Micheltore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.  
Lettuce and Sunflower ESTs from the Compositae Genome Project  
http://compenomics.ucdavis.edu/  
Unpublished  
Contact: Alexander Kozik [R.W.Micheltore]  
Department of Vegetable Crops, R.W.Micheltore Lab  
University of California at Davis (UCD)  
Asmundson Hall, UCD, Davis, CA 95616, USA  
Tel: 1-(530)-742-1742  
Fax: 1-(530)-752-9659  
Email: akozik@atgc.org [micheltore@vegmail.ucdavis.edu]  
belongs to contig QH\_CA\_Contig3643, see http://cgpdb.ucdavis.edu/  
for details.  
Plate: QH120 row: G column: 07.  
Location/Qualifiers  
1. .503  
/organism="Helianthus annuus"  
/mol\_type="mRNA"  
/cultivar="RHA801"  
/db\_xref="taxon:4232"  
/clone="QH120G07"  
/lab\_host="E.coli"  
/clone\_lib="QH ABCDI sunflower RHA801"  
/note="Vector: pBRCNAsf1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/  
TAG\_LIB=OH ABCDI sunflower RHA801  
TAG\_TISSUE=germinating seeds  
TAG\_SEQ=TCTGTGCGG"

FEATURES  
sourceBASE COUNT  
ORIGIN

Query Match 21.5%; Score 283; DB 13; Length 503;  
Best Local Similarity 72.9%; Pred. No. 4.8e-68;  
Matches 364; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db





Query Match 19.7%; Score 259.2; DB 14; Length 616;  
 Best Local Similarity 66.0%; Pred. No. 2.6e-61;  
 Matches 390; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

```

QY 292 CTTCCAGCGGGCCACTCTCTCGTGTGGCCCGCCGCGCAACTCTCCGACGTCGTGCG 351
D 29 CCACTCTCCTCATCTCCGCGCGCGCGCGCTCGCTGATCCCGTCGCCATCGT 88
QY 352 CGTCTCCATAGTCAACCCGGGCGCAATTCGCGCCCGCAGAGAGACTCTCCACCTACCC 411
D 89 CGTCACATCTAGTCAACCCCGCCAGTTCGCGCCCTCAGAGAGACTCTCCACCTACCC 148
QY 412 CTCGACTTCGAGCGGAGCTAAGAACTCGCGTCTGTTCCCGGGCGGTCAACGTCGT 471
D 149 TTCGACTTTCGCGTGACTCTCGCAAGCTCGCTCC--ACCGCGCTCGTGATGCCGT 205
QY 472 TTTCATPCCCGTAACTTGTACGATTACGGGAAGAGCGGTGTGTGAGTGCAGCAGGC 531
D 206 CTTCAACCCCGCTGACTCTACGTTCTGGAGCGGTGCTGCGGCGCGCTCCGAGG 265
QY 532 TGTGGAATGGTGTCTGCGTGTAGAGTGGTCCGGGCGAGAAAGTTGGGTGAGGTGA 591
D 266 CGCATCTCTCGTCTGGAGAGCGGTTCGGGATGGGCGACAGACGTGGGTTCGGGTGA 325
QY 592 GAAGCTGGAAATGGGCTGTGTGGAGAGCGAGCGCGCTTTCTTCAGAGGGGTGGCGAC 651
D 326 GCGATTGGAGAGGATGTGCGGGCGAGCGCTCCCGTGTCTTCGAGGGGTGGCCAC 385
QY 652 TGTGTGACGAAGTGTGTTAAATTTGTGGAGCGAGATGTGGTGTCTCGTAAAGAGA 711
D 386 CATAGTCTCAAGCTGTAAATCATCATCGAGCGGATGTGCTGTTCGGAGAGAGA 445
QY 712 TTATCAGAGTGGCGCTTATTCAGAGAGTGTTCGAGATCTTGATTTTCCATAAAGT 771
D 446 TTATCAGTGTAGTGGCGCTCATCTGCGTATGTTCTGATCTTGAATTTGCCATAGAT 505
QY 772 GATAGTGTGAAATTAACAGCTGATATGATGCGCTGGCAATGAGTTCAGTAATGTGA 831
D 506 AATAGATCAGAAATGTTCGAGAGAGCTGATGCTTTCGATGAGCTGTGGAATGTGA 565
QY 832 CTTTTCACCTGAAGAGAGGAAAGGCACTATCAATAAATAATCAATCTT 882
D 566 CATATCATCGAGGAGAGGAAAGGCATTTATCATAGTACTGCTGT 616

```

RESULT 12  
 AU290923  
 LOCUS AU290923 zinnia cultured mesophyll cell equalized cdna Zinnia  
 DEFINITION  
 eugens cdna clone Z5512, mRNA sequence.  
 ACCESSION  
 AU290923  
 AU290923.1 GI:24251431  
 VERSION  
 EST.  
 KEYWORDS  
 Zinnia  
 Zinnia elegans  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
 Heliantheae; Zinnia.  
 REFERENCE  
 1 (bases 1 to 545)  
 Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M.,  
 Matsuo, N., Minami, A., Nagata-Hiwatashi, M., Nakamura, K., Okamura,  
 Y., Sassa, N., Suzuki, S., Yazaki, J., Kikuchi, S. and Fukuda, H.  
 Visualization by comprehensive microarray analysis of gene  
 expression programs during transdifferentiation of mesophyll cells  
 into xylem cells  
 Proc. Natl. Acad. Sci. U.S.A. 99 (24), 15794-15799 (2002)  
 JOURNAL  
 COMMENT  
 Contact: Taku Demura  
 Morphogenesis Research Group  
 RIKEN Plant Science Center  
 1-7-22 Suehirocho, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9605

Fax: 81-45-503-9573  
 Email: demura@postman.riken.go.jp  
 This clone was obtained at our laboratory.  
 Seq primer: M13 forward.

## FEATURES

## Location/Qualifiers

## source

1..545  
 /organism="Zinnia elegans"  
 /mol\_type="mRNA"  
 /cultivar="Canary bird"  
 /db\_xref="taxon:34245"  
 /clone="Z5512"  
 /tissue\_type="mesophyll cell"  
 /clone\_lib="Zinnia cultured mesophyll cell equalized cDNA"  
 /note="Vector: pGEM-T easy; cultured in tracheary element  
 differentiation-inductive medium" 2 others  
 BASE COUNT 142 a 110 c 144 g 147 t  
 ORIGIN

Query Match 19.3%; Score 254.4; DB 9; Length 545;  
 Best Local Similarity 69.0%; Pred. No. 5.5e-60;  
 Matches 370; Conservative 0; Mismatches 148; Indels 18; Gaps 1;

```

QY 266 ATTGGCTGTCTCCACCATGGCTTCTCCACGCGGGCCACTCTCGTCTGCGCCAG 325
D 6 ATCGATTAGTCCCAACCATGGCTTCTCCACGCGGCACATCTCCCTCATCTGAA 65
QY 326 GCCGCCCACTCTCCGACGTCGTCGCCGTCTCCATCTACGTCAACCCGGGCCAATTCGCC 385
D 66 TCCCAAAACACACCACTCACCGTCTCAATCTAGTCAACCCCGTCAATCTCC 125
QY 386 CCCACGAGGACTCTCCACCTPACCCCTCCGACTTCGACGGCGAGTAAAGAACTTCGCG 445
D 126 CTCAACGAAGACTCTCCACATACCCATCTGACTTTAACGGCGACATTCAAAACCTCATC 185
QY 446 TCCGTTCCCGCGGCTGACGTCGTTTCATCCCGTAACTTGTACGATTCGAGGAG 505
D 186 TCCATCCCTAACCGCTTGATGTTGTTTCAACCCACATGATTTGTATGACTATGATGCT 245
QY 506 AACGCTGTGTGAGTGTGCGAGGCTGTGGAATGTTGCTGCGTGTGAGA----- 557
D 246 GTTGTGTAAACAGGGGAATGTTGAAACAGAGGGGTTGAGGAGAGAGAGAGGTG 305
QY 558 -----GTGGGTCCGGGCACGAAAGTTGGGTGAGGTTTGAGAGCTGGAATTTGGG 607
D 306 TCGTGTTCGAGGGTAAAGGGCATGAGACGTGGGTTAGGGTTGAGAGTTGAGAGAGG 365
QY 608 CTGCTGGGAAGACAGCGCCGCTTTCTTCAGAGGGGTGGGACTGTGGTGAAGAAGTTG 667
D 366 ATGTGTGGGAAAGTAGCCCGCTTTTTCAGAGGGGTGGCTACTGTTGTTGCTAAGTTG 425
QY 668 TTTAATATTGTGGAGCCAGATGTGCTGTGTTCCGTTAGAGAGGATATCAGCAGTGGCGG 727
D 426 TTTAATATTGCAACCCGATGTTGCGGTTTNGTAAAGAGGATATCAGCAGTGGCGT 485
QY 728 CTTATTCAGAGAGTGTTCGAGATCTTGATTTTTCATAAAAGTGATAGGTGCTGA 783
D 486 GTTATTCACGAGTGTTCGAGATCTAGATTTGGGTGTTAAGTGATCGGTTCTGA 541

```

## RESULT 13

## CB629808/c

## LOCUS

## DEFINITION

## CB629808

## OSIIEB06F09.r OSIIEB Oryza sativa (indica cultivar-group) cDNA

## clone OSIIEB06F09 3', mRNA sequence.

## ACCESSION

## CB629808

## VERSION

## CB629808.1

## KEYWORDS

## EST.

## SOURCE

## ORGANISM

## Oryza sativa (indica cultivar-group)

## Oryza sativa (indica cultivar-group)

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

## Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

## 1 (bases 1 to 746)





QY 564 CCGGGCAGAAAGTTGGTGAGGCTTGAGAGCTGGATTTGGGGCTGTGTGGAGAGCA 623  
 Db 249 ATCGGCACGAGACGTGGATTGGGTGGAGCGGCTGGAGAAAGGGGCTGTGTGGGAGCAGCA 190  
 QY 624 GGCCCGCTTTCTTTCAGAGGGGTGGCGACTGTGGTGACGAAAGTTGTTTAAATTTGTGGAGC 683  
 Db 189 GGCCTGTCTTCTCCGGCGGTGGCCACCGTGTGTGCGCAGCTGTTCAATATCGTGGAGC 130  
 QY 684 CAGATGTGGCTGTGTTCGGTGAAGAGATTATCAGCAGTGGCGGCTTTATTCAGAGGATGG 743  
 Db 129 CGGACGTGGCGGTGTTCGGGAAGAAGATTATCAGCAGTGGCGGCTCATTTGTGGGATGG 70  
 QY 744 T 744  
 Db 69 T 69

RESULT 15  
 BF064760  
 LOCUS  
 DEFINITION HV\_CB0017012f Hordeum vulgare seedling green leaf EST library  
 cDNA clone HV\_CB0017012f, mRNA sequence.  
 ACCESSION BF064760.1 GI:10841399  
 VERSION  
 KEYWORDS  
 SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
 ; Triticeae; Hordeum.  
 REFERENCE 1 (bases 1 to 838)  
 AUTHORS Wing,R., Close,I.J., Kleinohs,A., Wise,R., Wei,F., Begum,D.,  
 Frisch,D., Yu,X., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,  
 D.W., Fenton,R.D., Oates,R. and Main,D.  
 TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics: Blumeria infected incompatible (Mia6) seedling  
 leaf cDNA library  
 JOURNAL Unpublished  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Total hg bases = 449  
 Seq primer: AATTAACCTCTACTARAGGG  
 High quality sequence start: 9  
 High quality sequence stop: 692.

FEATURES  
 source  
 1..838  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="C116151 (Mia6)"  
 /db\_xref="taxon:112509"  
 /clone="HV\_CB0017012f"  
 /tissue\_type="seedling green leaf"  
 /lab\_host="SOLr"  
 /clone\_lib="Hordeum vulgare seedling green leaf EST  
 library HVcDNA0005 (Blumeria challenged)"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
 C.I. 16151 (Mia6) plants were greenhouse grown in the R  
 Wise lab at Iowa State University, Ames, IA; 7 day old  
 green seedlings were challenged with isolate 5874 (AvrMla6  
 ) of Blumeria graminis f. sp. hordei, and leaves were  
 harvested 20 and 24 hr post-inoculation and snap frozen;  
 uninoculated leaves were harvested 20 hr post-inoculation  
 (Wei, Wise). In the TJ Close lab at the University of  
 California, Riverside, total RNA was prepared from each  
 sample pool, equal quantities of all three RNA pools were  
 combined, poly(A) RNA was purified from the mixture, one

primary unamplified cDNA library was made, and 1 million  
 pfu were in vivo excised to give pBluescript SK(-) cDNA  
 phagemids (Choi, Close). Phagemids were plated and picked  
 at the Clemson University Genomics Institute (CUGI) (Begum  
 , Palmer, Frisch, Atkins and Wing). Plasmid DNA  
 preparations, DNA sequencing and sequence analysis were  
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates  
 , Rambo, Main). The sequence has been trimmed to remove  
 vector sequence and contains a minimum of 100 bases of  
 phred value 20 or above. For more details on library  
 preparation and sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders. Also  
 see Close TJ, Wing R, Kleinohs A, Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

BASE COUNT 188 a 208 c 240 g  
 ORIGIN

Query Match 18.5%; Score 244; DB 10; Length 838;  
 Best Local Similarity 62.5%; Pred. No. 5.7e-57;  
 Matches 436; Conservative 0; Mismatches 250; Indels 12; Gaps 3;  
 QY 400 CTCACCTACCCCTCGGACTTCGACGGCGACGTAAAGAAATCGCGTCCGTTCCCGCGG 459  
 Db 1 CTCCTCTACCCCTCTGACCTGGCCAGGGCGGTGCAGCAGCTCCCTAC---ACCGGCGC 57  
 QY 460 CGTCGACGTCGTTTTCCATCCCGTAACCTGTACGATTACGGGAAGAACGGTGTGGTGA 519  
 Db 58 GTCACACGCGGGGTGAACCCCAAAACCTCTACCAACCGCGCGCAGCTGTCTCTGCGCG 117  
 QY 520 CGTGGCAGAGGCTGTGTGGAATGTGTGCGCTTGAGAGTGTTCGGCGACGAAAGTTG 579  
 Db 118 CTCGCGCAGGCTCTCTCTTGCTG-----GAGCTGCGGGGACGGCAGCAACGTG 171  
 QY 580 GGTGAGGTTGAGAGCTGGAATGGGCTGTGTGGGAAGACAGCCGCTTTCTTCAG 639  
 Db 172 GATCCGGTGGAGCGGCTGGAGAAGGGCTGTGTGGGCCACCGCGCAGTCTTCTTCG 231  
 QY 640 AGGGTGGGACTGTGTGACGAAAGTTGTTTAAATTTGAGAGCAGATGTGGCTGTGT 699  
 Db 232 TGGGTGGCCACCGTCTGTCGCAAGCTGTTCAACGTCGTCGAAACCGCAGCTGCCATGTT 291  
 QY 700 CGGTAAAGAGGATTATCAGCAGTGGCGCTTATTCAGAGAGTGTTCGAGATCTTGATTT 759  
 Db 292 CGGCAAGAAGGATTATCAGCAATGCGCTCTCATCTGCCGAATGGTTCGACCTTGATTT 351  
 QY 760 TTCCATAAAGTGATAGGTGCTGAAATAACACGTGATAATGATGGCTGGCAATGATTC 819  
 Db 352 TGCCATAGAGATAGTAGGATCAGAAATAATGCGAGAAGCCGATGTTTGGCATGAGCTC 411  
 QY 820 AGTAAATGTCACCTTTACCTGAAGAGAGGAAAGGCACCTATCAATAAATAATCAAT 879  
 Db 412 TCGCAACGTGCACCTATCCCTGAAAGAAAGGAGAGGCGTTTCCATTAGTAGTCACT 471  
 QY 880 GTTAAGAGCTAAATCAGCAG---CAGGAGATGCTCAGGTGCATTTGTGAGACTTCACAA 936  
 Db 472 GGTAAATGCTAGAACTGCTGTGTTGAATAAGCAACAGTGTGCAAAACATATAAGGA 531  
 QY 937 TTGCTCATCCAAAGTGTACTGATGCTGGTGGAGAGGATCGATTATGTCGAGATGTTGA 996  
 Db 532 TCAGATAGTCAGACATTTTGGCTGAAGCTGGTGGTGGGTTGATTATGTTGAGATTGTGA 591  
 QY 997 TCAAAATAATTGGGAAGTGGACAGATCAAGAGTCTCTGTGCTTCTGTGTGTCTGC 1056  
 Db 592 GCAGAAAGTTTGGTACCTGGGGAGACGATTGACCGCCCTGTGTGTCATTGTGTGCCAC 651  
 QY 1057 ATGGTTTGGCAAGTCAGGCTTATAGACCAATGAAA 1094  
 Db 652 ATGGTTTGGAAAGGTTAATTGATCGCCCAATTTCGAA 689

Tue Dec 16 06:11:44 2003

us-10-033-269-8.rst

Page 12

Search completed: December 16, 2003, 00:07:42  
Job time : 3118 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 18:51:12 ; Search time 415 Seconds  
(without alignments)  
8560.146 Million cell updates/sec

Title: US-10-033-269-8

Perfect score: 1316

Sequence: 1 tcggcacagagcttctcag.....tggcattgctattgttagg 1316

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	47.5	1338	AAZ20926	Nucleotide sequenc
2	426.8	32.4	2184	AAC46075	Arabidopsis thalia
3	422.2	32.1	1189	AAC41869	Arabidopsis thalia
4	354	26.9	1264	AAZ20927	Nucleotide sequenc
5	140.8	10.7	908	AAC52194	Arabidopsis thalia
C 6	109	8.3	1672	AAI13694	Enterococcus faeca
C 7	109	8.3	1672	AE899489	Enterococcus faeca
C 8	101.2	7.7	10982	AA846264	DNA encoding novel

9	99.8	7.6	690	22	AAH53230	S. epidermidis ope
10	99.8	7.6	867	24	ABN92021	Staphylococcus epi
C 11	99.8	7.6	4103	22	AAH54816	S. epidermidis gen
C 12	99	7.5	4549	18	AAV74485	Staphylococcus aur
C 13	91.6	7.0	78845	21	AAH81483	N. meningitidis pa
C 14	91.6	7.0	349980	21	AAZ1608	Neisseria meningit
15	91.6	7.0	143768	21	AAH81490	N. meningitidis B
C 16	86.6	6.6	1132	22	AAH55061	S. meningitidis gen
17	83.6	6.4	546	24	ABK73744	Bacillus lichenifo
18	82.6	6.3	858	22	AA800989	B. subtilis panc g
19	82.6	6.3	2363	22	AA802314	B. subtilis pancBCD
20	82.6	6.3	4450	22	AA802327	B. subtilis panB
21	82.6	6.3	7140	24	AAH41228	Pantothenate relat
22	82.6	6.3	7140	24	AAH48884	Plasmid pAN441 for
C 23	82.6	6.3	7381	22	AA802335	B subtilis panB ge
C 24	82.6	6.3	8503	22	AA802334	Plasmid pAN006 car
C 25	82.6	6.3	10801	22	AA802323	Plasmid pAN004 car
26	79.8	6.1	858	24	ABQ90275	M. capsulatus gene
27	79.2	6.0	777	25	ABZ39410	N. gonorrhoeae nuc
28	78	5.9	1193	25	ABA00659	Human ENZM-7 cDNA,
29	78	5.9	1446	23	AA892821	DNA encoding novel
C 30	73.4	5.6	640681	24	ABJ92787	Buchnera sp. genom
31	71.6	5.4	100848	22	AAZ28552	Genomic fragment #
32	70.8	5.4	837	22	AAH65097	C glutamicum codin
33	70.8	5.4	960	22	AAH72059	Corynebacterium gl
34	70.8	5.4	2164	21	AAA71997	C. glutamicum panB
35	70.8	5.4	2164	21	AAA40273	C. glutamicum panB
36	70.8	5.4	2164	21	AAA40282	C. glutamicum panB
C 37	70.8	5.4	349980	22	AAH64966	C glutamicum codin
C 38	69.2	5.3	4403765	22	AAI99683	Mycobacterium tube
C 39	69.2	5.3	4411529	22	AAI99682	Mycobacterium tube
C 40	67.4	5.1	837	23	AA888471	DNA encoding novel
C 41	67.4	5.1	2712	23	AA886723	DNA encoding novel
C 42	67.4	5.1	2712	23	AA889038	DNA encoding novel
C 43	67.4	5.1	2712	23	AA889081	DNA encoding novel
44	67.4	5.1	2712	23	AA890054	DNA encoding novel
45	67.4	5.1	3061	23	AA889720	DNA encoding novel

#### ALIGNMENTS

##### RESULT 1

AAZ20926

ID AAZ20926 standard; DNA; 1338 BP.

AC AAZ20926;

DT 02-DEC-1999 (first entry)

XX Nucleotide sequence of L. japonicus Pantothenate Synthetase (PS).

XX pantothenate synthetase; cloning; herbicide; biosynthesis;

XX Coenzyme A precursor; pantoate-beta-alanine ligase; ds.

XX Lotus japonicus.

XX Key

XX Location/Qualifiers

XX CDS

XX /tag= a

XX /product= Pantothenate Synthetase

XX WO9942565-A1.

XX 26-AUG-1999.

XX 02-JUN-1998;

XX 98WO-EP03261.

XX 31-MAY-1997;

XX 97GB-0011163.

XX 27-JUN-1997;

XX 97GB-0013477.

XX (AGRE ) HOECHST-SCHERING AGREVO GMBH.

XX

PI Abell C, Smith AG, Genschel U, Laber B;  
 XX WPI; 1999-527466/44.  
 DR P-PSDB; AAY42402.  
 XX  
 PT New isolated DNA molecule, useful in herbicidal compositions -  
 XX  
 PS Claim 4; Fig 1.2; 64pp; English.  
 XX  
 CC This is the nucleotide sequence of the Lotus japonicus Pantothenate  
 CC synthetase (PS) enzyme, which is essential in the biosynthesis of  
 CC Coenzyme A.  
 CC The invention provides methods to assay proteins for PS activity, and  
 CC the ability of compounds to inhibit PS.  
 CC The ability of compounds which are found to inhibit PS can be used as  
 CC herbicides.  
 CC The nucleotide probe can be used to amplify PS coding sequences from a  
 CC chosen organism via the process of polymerase chain reaction.  
 CC A number of assays have previously been used to measure PS activity  
 CC but are unsuitable for large screen biochemical screening of compounds  
 CC to find PS inhibitors. In this invention however the isolation of the  
 CC DNA molecule encoding PS enables a host cell to produce the protein  
 CC which can then be tested against compounds potentially able to inhibit  
 CC the enzyme.  
 XX  
 SQ Sequence 1338 BP; 380 A; 242 C; 328 G; 388 T; 0 other;

Query Match 47.5%; Score 625; DB 20; Length 1338;  
 Best Local Similarity 79.0%; Pred. No. 1.5e-155;  
 Matches 800; Conservative 0; Mismatches 195; Indels 18; Gaps 4;

QY 194 GCCCAAGGGTGATCTCGACAGGCTCGATCGGAGCTGTGCGCTCGATGGGGCC 253  
 DB 23 GCACCAATGGTGATATCTGATAAGGACGAGATGCGGAAATGGTCAAGTCCATGCGATCC 82  
 QY 254 CAGGCAAGCTCATTTGGGTGCTCCACCTCCAGCTGGGCTTCCACCGGGCCACTCTCG 313  
 DB 83 CAAGCAAGCTCATCGGCTGTTCCACCATGGGCTTCCCTCAGAGGGCCACTTCT 142  
 QY 314 CTCGTGGCCCGGCGGCAACTCTCCGACGTGCTGCGGCTCTCCATCTAGTCAACCGC 373  
 DB 143 CTCGTGACAGAGCGTCAACACCGCTGACCTGCTGCGGCTCTCAATCTATGTCAACCT 202  
 QY 374 GGCATATGCGCCCGACGAGGACCTCTCCACCTACCCCTCGACTTCGACGGGACGTA 433  
 DB 203 GGCCAGTTTCCCGACGAGGACCTTTCGACATACCTTCTGATTTTCAAGGTGATCTC 262  
 QY 434 AAGAACTCGGTGCTGTTCCCGGGCGTGCAGCTGCTGTTTCCATCCCGTAATCTGTAC 493  
 DB 263 CAAAACTCATGCTGTTCTGTTGTTGTTGTTGTTTCCACCCCAATTTGAT 322  
 QY 494 GATTACGGGAAGAACGGTGGTGGTGAACGTGGCAGAG-----GCTGGTGAATGGTGTG 547  
 DB 323 GATTACGGTGGTGGTGGGTTGATGCTGTCGCGAGTGTGTTGATGGGTGGTGTCT 382  
 QY 548 TGCCTTGA-----GAGTGGGTCCGGACGAAAGTGGGTGAGGGTTGAGAACTGAA 601  
 DB 383 TGTGTTGATAGGAGGAGTGGTTTGGGCATGAAACTTGGGTAGAGCTGAGAACTGGAG 442  
 QY 602 TTGGGGTGTGTGGGAAGAGCAGCCCGTTTCTTCAGAGGGTGGCGACTGTGTTGACG 661  
 DB 443 AAACCCCTTGTGGGAAGAGTAGGCCCTGTTTCTTTAGGGGGTGGCCACCATTTGTTACC 502  
 QY 662 AAGTTGTTTAATTGTTGGAGCCAGATGTGCTGTGTTGGTGAAGAAGGATTATCAGCAG 721  
 DB 503 AAGTTGTTTAATTGTTGGAGCCATGATGTTCTGTTTGGGAAGAGGACTATCAGCAA 562  
 QY 722 TGGCGGCTTATTCAGAGATGGTTCGAGATCTGATTTTCCATAAAGTATAGGTGCT 781  
 DB 563 TGGAAATTAATTCAGAGATGGTTCGAGATCTGATTTTCCATAAAGTATAGGTGCT 622  
 QY 782 GAAATAACAGTGAATATGATGGCTTGGCATGATTCAGTATGTGACCTTTCACT 841

DB 623 GAAGTAATACGTGAGAAAGATGGCCTAGCAATAGTTCCTCCGTAATGTGCTACTACCT 682  
 QY 842 GAAGAGAGGAAAGGCACTATCAATTAATAAATCAATTTTAAGAGCTAAATCAGCACA 901  
 DB 683 GAAGAGAGGAAAGGCGAGTATCTATAAATAAATCAATTTTAGAGCTAAATTCGCGACA 742  
 QY 902 GGAGATGCTCAGGTGCATTTGTGAGAAAGTTGACAAATTTGGTCATCCAAAGTGTACTGAT 961  
 DB 743 GAAGATGACAGATACATTTGTGAGAAATTTGATAAATTTGGTCGTCGCAAGTATCACCGA 802  
 QY 962 GCTGGTGAAGGATCGATTAATGCTGAGATTTGATCAAAATAAATTTGGAGAAAGTGAA 1021  
 DB 803 GCTGGTGAAGGATTTGATTAATGCTGAGATTTGATCAAAATAAATTTGGAGAAAGTGAA 862  
 QY 1022 CAGATCAAGAGTCCCTGCTCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081  
 DB 863 TGGATCAAGGTCCTGTTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 922  
 QY 1082 GACAACTGGAATCAACTTGTCAATGAATGT---TTGATCTAACCTTCTG---TCATCT 1135  
 DB 923 GACAACTAGAAATCAACTTGTAAATGGAAGTAAAGATGATCTAACCTTGTGAATAATCT 982  
 QY 1136 CAACATGGGCGCACATGCTTAATTAATAGTTCGGGCGCACGCTTAACAATTC 1188  
 DB 983 CAGACATGGACCATATGATTAGTAGTCTGTCATTTTCATGGGGTATAGACTTC 1035

RESULT 2  
 AAC46075  
 ID AAC46075 standard; DNA; 2184 BP.  
 AC AAC46075;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48823.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 OS Arabidopsis thaliana.  
 XX  
 EP1033405-A2.  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135153.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136352.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140553.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.



PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 32.1%; Score 422.2; DB 21; Length 1189;  
Best Local Similarity 66.8%; Pred. No. 9.9e-102;

Matches 623; Conservative 0; Mismatches 298; Indels 12; Gaps 1;

QY 197 CCAAGGGTGATCTCCGACAGGCCTCGATCGGAGCTGGTCTGATCGGCGCCAG 256









PN WO98050555-A2.  
 XX 12-NOV-1998.  
 PD  
 XX 04-MAY-1998; 98WO-US08985.  
 XX  
 PR 14-NOV-1997; 97US-0066009.  
 PR 06-MAY-1997; 97US-0044031.  
 PR 16-MAY-1997; 97US-0046655.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Barash SC, Dillon PJ, Kunsch CA;  
 PI WPI; 1999-045171/04.  
 DR  
 XX  
 XX New isolated Enterococcus faecalis polynucleotides and polypeptides  
 PT - used to develop products for the detection of Enterococcus and for  
 PT use in vaccines for prevention or attenuation of Enterococcus  
 PT infection.  
 XX  
 PS Claim 1; Page 1970; 2084pp; English.  
 PS  
 XX A computer readable medium has been developed which has recorded on it  
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.  
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are  
 CC primary nucleotide sequences, also known as contigs. The computer-based  
 CC system can identify fragments of the Enterococcus faecalis genome with  
 CC commercial importance. The products can be used to detect the presence  
 CC of Enterococcus faecalis in samples. They can also be used for  
 CC diagnosing Enterococcal infection in an animal and monitoring  
 CC progression of disease, and for identifying agents which can be used to  
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or  
 CC another related organism, in vivo or in vitro. In particular the  
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences  
 CC can be used in vaccines to prevent or attenuate an Enterococcal  
 CC infection.  
 XX  
 SQ Sequence 1672 BP; 487 A; 365 C; 173 G; 634 T; 13 other;  
 PI  
 XX  
 XX Query Match  
 XX Best Local Similarity 8.3%; Score 109; DB 20; Length 1672;  
 XX Matches 275; Conservative 2; Mismatches 258; Indels 3; Gaps 1;  
 QY 605 GGGCTGTGTGGGAAGACAGAGCGCCGCTTTCTTCAGAGGGGTGGCGACTGTGGTGCAGAG 664  
 Db  
 QY 593 GGTATTATGTGGAGCTAGTAGGCCAATACATTTTAGAGGGGTTCACCTGTTAGTAAR 534  
 Db  
 QY 665 TTGTTTATATGTGGAGCCAGATGTGGCTGTGTTGGTAAAGAGGATATACAGCAGTGG 724  
 Db  
 QY 533 TTATTTAAATATACAGCAGATAGAGCTTATTTGGGAGAAAGGATGCACACAGTTA 474  
 QY 725 CGGCTTATTCAGAGATGTTTCGAGATCTTGATTTTCCATAAAGTGNATAGTCTGAA 784  
 Db  
 QY 473 GCTGTAATTAAGAATGTTAGGACCTTAATTTGATATATAGACGTTGTAGGGGTGCTCA 414  
 QY 785 AFAACACGTGTAATGATGGCTGGCAATGAGTTCACGTAATGTGCACCTTTCACTGAA 844  
 Db  
 QY 413 ATTATAAGGAAGAGATGAGTACGTAAAGTCTAGAAACACATCTTAAGTTAGAG 354  
 QY 845 GAGAGGGAAGGCTATCAATAAATAATCATTTGTTAAAGCTAAATCAGCAGCAGGA 904  
 Db  
 QY 353 GAAGATCTTCAGCTACTATTTTAAATAATCTTTAACTTTAGCYAAGGAGCCCTTAAT 294  
 QY 905 GATGTCAGTGCATGTGAGAGATGACAAATTTGGTGCATCCAAAGTGT---TACTGAT 961  
 Db  
 QY 293 AATGAGAGAGATAGCTGTAAATAATAATKAGATAATAAGTAAATAATAATAACTTGT 234  
 QY 962 GCTGTGGAAGGATCGATTATCTGAGATTTGATCAAAATAATTTGGAGAAAGTGGAA 1021  
 Db  
 QY 233 AATTTGGCAAGATTTGATTAATGAGATTTGTTGATCTTTATCATTAACAAGATAAT 174  
 QY 1022 CAGATCAAGAGTCTGTCTCTTCTGTTGTTGCTGATGTTTGGCAAGTCAGGCTTATA 1081

Db 173 TATATTGAAAAATCAGTTTTAGTTGCAATAGCAGTATTTATAGAAAAACAAGATTATA 114  
 QY 1082 GACAACATGGAATCAACTTGCTCAATGAATGTTGATCTAAACCTTCTGTCATCTCAA 1139  
 Db 113 GATAATTTTACATTTGAATTACAATAGGGGCTATAATAATGANAGTTCATGTTAAA 56  
 RESULT 7  
 ABS99489/c  
 ID ABS99489 standard; DNA; 1672 BP.  
 XX  
 XX AC ABS99489;  
 XX  
 XX 18-DEC-2002 (first entry)  
 XX  
 XX Enterococcus faecalis contig sequence #757.  
 XX  
 XX Computer readable medium; Enterococcus faecalis; microbe; growth;  
 KW pathogenicity; vaccine; resistance; Enterococcal infection; commercial;  
 KW therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;  
 KW biotech technology; antibacterial; modulator of nucleic acid expression;  
 KW contig; ds.  
 XX  
 XX Enterococcus faecalis.  
 OS  
 XX US2002120116-A1.  
 XX  
 XX 29-AUG-2002.  
 XX  
 XX 04-MAY-1998; 98US-0070927.  
 XX  
 XX 04-MAY-1998; 98US-0070927.  
 XX  
 XX (KUNS/) KUNSCH C A.  
 PA (DILL/) DILLON P J.  
 PA (BARA/) BARASH S.  
 XX  
 XX Kunsch CA, Dillon PJ, Barash S;  
 XX  
 XX WPI; 2002-750065/81.  
 DR  
 XX  
 XX Computer readable medium having recorded on it a Enterococcus faecalis  
 PT nucleotide sequence useful for detecting diseases related to  
 PT Enterococcus infections in animals  
 XX  
 XX Claim 1; Page -; 119pp; English.  
 PS  
 XX The present invention relates to a new computer readable medium with an  
 CC Enterococcus faecalis nucleotide sequence. The invention is useful to  
 CC diagnose the presence of E.faecalis in a sample or determining the  
 CC presence of a specific microbe in a sample. The invention is also useful  
 CC for modulating the growth or pathogenicity of E.faecalis, in a vaccine  
 CC to confer resistance to Enterococcal infection, for commercial,  
 CC therapeutic and industrial purposes, and for fermenting a particular  
 CC sugar source or to produce a particular metabolite. The invention is  
 CC useful for detecting diseases related to Enterococcus infections in  
 CC animals, and for detecting E.faecalis using biotech technology. The  
 CC present nucleic acid sequence represents an Enterococcus faecalis contig  
 CC DNA sequence of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification but was obtained in electronic format directly  
 CC from USPTO at <http://seqdata.uspto.gov>.  
 XX  
 SQ Sequence 1672 BP; 487 A; 365 C; 173 G; 634 T; 13 other;  
 Query Match  
 Best Local Similarity 8.3%; Score 109; DB 24; Length 1672;  
 Matches 275; Conservative 2; Mismatches 258; Indels 3; Gaps 1;  
 QY 605 GGGCTGTGTGGGAAGAGCAGCGCCGCTTTCTTCAGAGGGGTGGCGACTGTGGTGCAGAG 664  
 Db  
 QY 593 GGTATTATGTGGAGCTAGTAGGCCAATACATTTTAGAGGGGTTCACCTGTTAGTAAR 534











CC Neisserial bacteria. For example, some of the identified proteins could  
 CC be components of vaccines against *Meningococcus B*; against all serotypes;  
 CC and/or against all pathogenic *Neisseria*. Identification of sequences  
 CC from the bacterium will also facilitate production of biological probes,  
 CC particularly organism-specific probes. Attempts to make efficacious  
 CC *Meningococcus B* vaccines have failed mainly due to antigen tolerance.  
 CC Multivalent vaccines have also been tried but none have successfully  
 CC overcome antigenic variability. The provision of further, complete  
 CC sequences may provide an opportunity to identify secreted or surface  
 CC exposed proteins that may be presumed targets for the immune system and  
 CC which are not antigenically variable or at least more conserved than  
 CC other more variable regions.

XX SQ Sequence 78845 BP; 19372 A; 21111 C; 19042 G; 19318 T; 2 other;

Query Match 7.0%; Score 91.6; DB 21; Length 78845;  
 Best Local Similarity 48.8%; Pred. No. 4.7e-13;  
 Matches 247; Conservative 0; Mismatches 259; Indels 0; Gaps 0;  
 QY 593 AAGCTGGAATTGGGCTGTGGGAAGAGAGCGGCCGTTTCTTCAGAGGGGTGGCGACT 652  
 Db 3833 AATCTGCAAAATGAGTTGTGCGCAAAATTCGCCCGGGGCATTTCGCGGTGGCAACG 3774  
 QY 653 GTGGTGACAGATTGTTAAATTTGTGGAGCCAGATGGCTGTGTTTCGTTAAGAGAT 712  
 Db 3773 GTTGTCTTAAATTTCCACATCGTTTCCCGGACATTGCTGTGTTGTTAAGAGAT 3714  
 QY 713 TATCAGCAGTGGCGCTTATTCAGAGGATGGTTTCGAGATCTTGAATTTTCCATAAAGTG 772  
 Db 3713 TACCAGAGCTTGGCGTGAATTAAGTTTGTGCAAGATTTGAATTTGATGTTGAATA 3654  
 QY 773 ATAGGTGCTGMAATAACACGTGATAATGATGGCTGGCAATGAGTTCACGTAATGTGCAC 832  
 Db 3653 GTGGCTGTGTATACAGGGCGCGGAGACGCGGTGGCACTGTGAGCGCGCAACGATAT 3594  
 QY 833 CTTTCACCTGAGAGAGGAAAGGCACATATCAATAAATAATCATTTGTTAAGAGCTAAA 892  
 Db 3593 TTGAGTGGCGGGAACGCGAGAGACCGCGCTGTACCGCGAATTAAGGCTGTTGCC 3534  
 QY 893 TCACAGCAGGAGATGGTCAGGTGATGTCAGAGATTGACAAATTTGTCATCCAAAGT 952  
 Db 3533 GAATTCCTTGGTGAGGCGAGTTGGATTATCAGGTTTGGAAACGTCGCGCTCCATCC 3474  
 QY 953 GTTACTGATGCTGGTGAAGATCGATTATGCTGAGATTGTTGATCAAAATAAATTGGAG 1012  
 Db 3473 CTGACAGAAATACGCTGGTGGTGCATATGTCGAATTCGGAATCCCGCGCGGATACGCTGAA 3414  
 QY 1013 AAGTGGAACAGATCAGAGTCCTGCTCTTCTGTGTTGTCATGTTGGCAAGTC 1072  
 Db 3413 GTGGCGCGGCGGAGATAGAAACTGGTGTCTTGGCGCGCGCTGTCTGGGACGACG 3354  
 QY 1073 AGGCTTATAGACAACATGGAATCAA 1098  
 Db 3353 CGCTGATTGCAATTTGGAATAAA 3328

RESULT 14

ID AAF21608

XX AAF21608 standard; DNA; 349980 BP.

XX AAP21608;

XX 13-MAR-2001 (first entry)

DE *Neisseria meningitidis B* nucleotide sequence SEQ ID NO:109.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;  
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
 KW ds.

OS *Neisseria meningitidis*.

XX WO200066791-A1.

XX 09-NOV-2000.  
 XX 08-MAR-2000; 2000WO-US05928.  
 XX 30-APR-1999; 99US-0132068.  
 PR 08-OCT-1999; 99WO-US23573.  
 PR 28-FEB-2000; 2000GB-0004695.  
 XX (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;  
 PI Galeotti C, Moia M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;  
 PI Frazer CM, Grandi G;  
 XX WPI; 2000-647603/62.  
 DR *Neisseria meningitidis B* full length genome sequence and open reading  
 PT frames are used to detect, treat and prevent *Neisseria* infections -  
 XX Claim 7; Appendix A; 592pp; English.  
 XX The present invention describes the full length genome of  
 CC *Neisseria meningitidis B* (NMB). The sequences in AAF21544 and AAF21607  
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the  
 CC sequence was too long to go in a record on its own it was split into 8  
 CC sequences which overlap each other at the beginning and end of each  
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at  
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 is repeated at  
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the  
 CC *Neisseria* proteins given in AAB58550 to AAB58593, and AAF21589 to  
 CC AAF21606 represent PCR primers which are used in the exemplification of  
 CC the present invention. The NMB genome and fragments from it have  
 CC antibacterial activity, and can be used in vaccines and gene therapy.  
 CC *Neisseria* nucleic acids, proteins and/or antibodies which binds to the  
 CC proteins can be used in compositions for treating or preventing infection  
 CC due to *Neisseria* bacteria or as a diagnostic reagent for detecting the  
 CC presence of *Neisseria* bacteria or of antibodies raised to *Neisseria*  
 CC bacteria. Computers, computer memory, computer storage medium or computer  
 CC databases can be used in a search to identify open reading frames (ORFs)  
 CC or coding sequences within the NMB genome. The DNA sequences provide  
 CC further opportunities to find antigenic or immunogenic proteins which are  
 CC more effective in vaccines than the outer membrane proteins currently  
 CC used.

XX SQ Sequence 349980 BP; 82523 A; 82940 C; 96712 G; 87805 T; 0 other;  
 Query Match 7.0%; Score 91.6; DB 21; Length 349980;  
 Best Local Similarity 48.8%; Pred. No. 9.6e-13;  
 Matches 247; Conservative 0; Mismatches 259; Indels 0; Gaps 0;  
 QY 593 AAGCTGGAATTGGGCTGTGGGAAGAGAGAGCGGCCGTTTCTTCAGAGGGGTGGCGACT 652  
 Db 290737 AATCTGCAAAATGAGTTGTGCGGCAAAATTCGCCCGGGGCATTTCGCGGTGGCAACG 290756  
 QY 653 GTGGTGACAGATTGTTAAATTTGTGGAGCCAGATGTCGCTGTTCGTTAAGAGAT 712  
 Db 290797 GTTGTCTTAAATTTGTTCACATCGTTTCCCGGACATTGCTCTGTTGTTAAGAGAT 290856  
 QY 713 TATCAGCAGTGGCGCTTATTCAGAGGATGGTTTCGAGATCTTGAATTTTCCATAAAGTG 772  
 Db 290857 TACCAGCAGCTTGGCGTGAATTAAGTTTGTGCAAGATTTGAATTTGATGTTGAATA 290916  
 QY 773 ATAGGTGCTGAAATAACACGTGATATGATGGCTGGCAATGAGTTCACTAATGTGCAC 832  
 Db 290917 GTCCCTGTTGATACAGGCGCGCGAAGACGCGTTGGCACTGTCGAGCGCGCAACGATAT 290976  
 QY 833 CTTTCACCTGAGAGAGGGAAGGCACATATCAATAAATAAATCATTTGTTAAGAGCTAAA 892  
 Db 290977 TTGAGTGGCGGGAACGCGACGAGCAGCGCGCTGTACCGCGAATTAAGGCTGTTGCC 291036  
 QY 893 TCAGCAGCAGGAGATGGTTCAGGTGCATTTGTGAGAGTTTGACAAATTTGTCATCCAAAGT 952



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 23:15:54 ; Search time 471 Seconds  
(without alignments)  
9286.320 Million cell updates/sec

Title: US-10-033-269-8

Perfect score: 1316

Sequence: 1 tcggcagggctttctcag.....tggcattgtattgttagg 1316

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1316	100.0	1316	13	US-10-033-269-8
2	506	38.4	1183	13	US-10-033-269-6
3	389.8	29.6	1831	13	US-10-033-269-3
4	384.4	29.2	1235	13	US-10-033-269-12
5	382.6	29.1	1443	13	US-10-033-269-1
6	346.2	26.3	1148	13	US-10-033-269-10
7	300.2	22.8	1644	13	US-10-033-269-5
8	118.8	9.0	9381	15	US-10-114-170-7
9	109	8.3	1672	10	US-09-070-927A-757
10	99	7.5	4549	8	US-08-781-986A-174
11	83.6	6.4	546	10	US-09-974-300-1035
12	73.4	5.6	640681	10	US-09-790-988-1
13	70.8	5.4	837	10	US-09-738-626-132
14	70.8	5.4	3309400	10	US-09-738-626-1
15	67.2	5.1	501	10	US-09-974-300-5461

16	67.2	5.1	501	10	US-09-974-300-5490
17	57.8	4.4	999	15	US-10-156-761-4671
18	57.8	4.4	9025608	15	US-10-156-761-1
19	52.8	4.0	520	15	US-10-184-644-332
20	52.8	4.0	520	15	US-10-184-644-332
21	52.6	4.0	1167	15	US-10-184-634-332
22	52.6	4.0	9025608	15	US-10-156-761-2743
23	51	3.9	390	9	US-10-156-761-1
24	49.4	3.8	440	13	US-09-790-399-7
25	49.4	3.8	440	13	US-10-063-685-52
26	49.4	3.8	440	15	US-10-184-644-202
27	49.2	3.7	1392	15	US-10-184-634-202
28	49	3.7	390	9	US-10-156-761-3990
29	48.8	3.7	825	15	US-09-790-399-7
30	48.8	3.7	2274	15	US-10-156-761-2303
31	48.4	3.7	543	15	US-10-156-761-5313
32	48.2	3.7	1416	15	US-10-156-761-7012
33	48	3.6	1209	15	US-10-156-761-6191
34	48	3.6	1479	15	US-10-156-761-6772
35	47.8	3.6	4863	9	US-10-156-761-4461
36	47	3.6	1475	13	US-09-815-242-4071
37	46.8	3.6	939	15	US-10-271-889-12
38	46.8	3.6	1428	13	US-10-156-761-7243
39	46.8	3.6	1429	13	US-10-289-757-149
40	46.8	3.6	1440	13	US-10-289-757-40
41	46.8	3.6	1749	15	US-10-289-757-43
42	46.6	3.5	11058	15	US-10-100-679-58
43	46.4	3.5	4674	13	US-10-156-761-3629
44	46.4	3.5	4737	13	US-09-940-316B-26
45	46.2	3.5	1440	15	US-09-940-316B-30
					US-10-156-761-6190

#### ALIGNMENTS

#### RESULT 1

US-10-033-269-8  
; Sequence 8, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase  
; FILE REFERENCE: B01446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 1316  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-033-269-8

Query Match 100.0%; Score 1316; DB 13; Length 1316;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCGGCAGGAGGCTTCTCAGCAGGAGGCTCGTGGCGGTACACACCAAAACGTACTTCAAC 60  
Db 1 TCGGCAGGAGGCTTCTCAGCAGGAGGCTCGTGGCGGTACACACCAAAACGTACTTCAAC 60  
QY 61 CCAAGAGTCATAATAAATTTTGCCTTAACTCTTCAACACCGGTTCCGGCCAT 120  
Db 61 CCAAGAGTCATAATAAATTTTGCCTTAACTCTTCAACACCGGTTCCGGCCAT 120  
QY 121 CAGGTCGCTTCTTCACTTATTACACTCTTCAAGAAATCCCACTGTTTCAATATTTT 180  
Db 121 CAGGTCGCTTCTTCACTTATTACACTCTTCAAGAAATCCCACTGTTTCAATATTTT 180

QY 181 CA CAATGGCTCAGGCCCAAGGGTGAATCTCCGACAAAGGCTCGATCGGAGCTGTCGCG 240  
Db 181 CA CAATGGCTCAGGCCCAAGGGTGAATCTCCGACAAAGGCTCGATCGGAGCTGTCGCG 240  
QY 241 CTCGATGCGGGCCAGGGCAAGCTCATTTGGGCTGTCCTCCACCATGGGCTTCTCCAGCG 300  
Db 241 CTCGATGCGGGCCAGGGCAAGCTCATTTGGGCTGTCCTCCACCATGGGCTTCTCCAGCG 300  
QY 301 GGGCCACCTCTCGCTGCTGGCCAGGCCCGCCCAACTCTCCGACGTCGTCGCCGCTCTCCAT 360  
Db 301 GGGCCACCTCTCGCTGCTGGCCAGGCCCGCCCAACTCTCCGACGTCGTCGCCGCTCTCCAT 360  
QY 361 CTACGTCAACCGGGCAATTCGCGCCACGAGGACCTCTCCACCTACCCCTCCGACTT 420  
Db 361 CTACGTCAACCGGGCAATTCGCGCCACGAGGACCTCTCCACCTACCCCTCCGACTT 420  
QY 421 CGACGGCAGCTAAAGAACTCGCTGCTCCGCGGCGGCGCTGCGAGCTGTTTCCATCC 480  
Db 421 CGACGGCAGCTAAAGAACTCGCTGCTCCGCGGCGGCGCTGCGAGCTGTTTCCATCC 480  
QY 481 CGGTAACTTGTAACGTAACGGGAAGAACGGTGGTGGTGACGTCGGCAGAGGCTGGTGAAT 540  
Db 481 CGGTAACTTGTAACGTAACGGGAAGAACGGTGGTGGTGACGTCGGCAGAGGCTGGTGAAT 540  
QY 541 GGTGTCGCTGCTGAGAGTGGTCCGGGCAAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
Db 541 GGTGTCGCTGCTGAGAGTGGTCCGGGCAAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGT 600  
QY 601 ATTGGGGCTGTGCGGAAGAGCAGGCCCGCTTTCTTCAGAGGGTGGGAGCTGTTGGTAC 660  
Db 601 ATTGGGGCTGTGCGGAAGAGCAGGCCCGCTTTCTTCAGAGGGTGGGAGCTGTTGGTAC 660  
QY 661 GAAGTTGTTTAAATTTGGAGCCAGATGTGGCTGTGTTTGGTGAAGAGGATTAACGCA 720  
Db 661 GAAGTTGTTTAAATTTGGAGCCAGATGTGGCTGTGTTTGGTGAAGAGGATTAACGCA 720  
QY 721 GTGGGGCTTATTCAGAGGATGTTGAGATCTTGATTTTCCATAAAGTGAAGTGGTGG 780  
Db 721 GTGGGGCTTATTCAGAGGATGTTGAGATCTTGATTTTCCATAAAGTGAAGTGGTGG 780  
QY 781 TGAATAACACGTGATGATGCGCTCGCAATGAGTTCAGTAATGTCACCTTTCACC 840  
Db 781 TGAATAACACGTGATGATGCGCTCGCAATGAGTTCAGTAATGTCACCTTTCACC 840  
QY 841 TGAAGAGGGAAGGACCTATCAATAAATAAATCAATTTGAAGAGCTAATCAGCAGC 900  
Db 841 TGAAGAGGGAAGGACCTATCAATAAATAAATCAATTTGAAGAGCTAATCAGCAGC 900  
QY 901 AGGAGATGCTCAGTGCAATTCGAGAGTTCACAAATTTGGTCAATCCAAAGTGTACTGA 960  
Db 901 AGGAGATGCTCAGTGCAATTCGAGAGTTCACAAATTTGGTCAATCCAAAGTGTACTGA 960  
QY 961 TGCTGCTGGAAGGATCGAATATGCTGATGTTGATCAAAATTAATTTGGAGAAAGTGA 1020  
Db 961 TGCTGCTGGAAGGATCGAATATGCTGATGTTGATCAAAATTAATTTGGAGAAAGTGA 1020  
QY 1021 ACAGATCAAGAGTCTGCTGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 1021 ACAGATCAAGAGTCTGCTGCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1081 AGACAACTGAATCAACTTGTCAATGAATGTTGATCTAACCTTCTGCTCATCTCAAC 1140  
Db 1081 AGACAACTGAATCAACTTGTCAATGAATGTTGATCTAACCTTCTGCTCATCTCAAC 1140  
QY 1141 ATGGGCCACATGCTTAATTAATAGTTTGGGCCACGTCGTTTAACTTCTACAGTTTCATG 1200  
Db 1141 ATGGGCCACATGCTTAATTAATAGTTTGGGCCACGTCGTTTAACTTCTACAGTTTCATG 1200  
QY 1201 GTTATAGTCATGACAAATTTTTTTTCTGCCAGCATACATGATTAATTTGATGATGATTT 1260  
Db 1201 GTTATAGTCATGACAAATTTTTTTTCTGCCAGCATACATGATTAATTTGATGATGATTT 1260  
QY 1261 TACCGCATCAAAATTTCTATGAGAGCTGCTACAGTGGCATTTGCTATTGCTAGG 1316

Db 1261 TACCGCATCAAAATTTCTATGAGAGCTGCTACAGTGGCATTTGCTATTGCTAGG 1316  
RESULT 2  
US-10-033-269-6  
; Sequence 6, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harwell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase  
; FILE REFERENCE: B81446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
; US-10-033-269-6  
Query Match 38.4%; Score 506; DB 13; Length 1183;  
Best Local Similarity 72.1%; Pred. No. 6.4e-142;  
Matches 691; Conservative 0; Mismatches 255; Indels 12; Gaps 2;  
QY 197 CCAAGGCTGATCTCCGACAAAGGCTCGATGCGAGCTGGTGGCTCGATGCGGGCCCGAG 256  
Db 144 CCGCTGCTCATCACCGACAGGCGAGATCGGGAGTGGTGGAGTCCGCGGCGCGCAG 203  
QY 257 GCGAAGCTCATTTGGGCTGGTCCCCACCATGGGCTTCTCCACGCGGCCACCTCTCGCTC 316  
Db 204 GCGAGGACCATGCGCTTCCGCGCCACCATGGGTTTCTCCACGAGGCGCACCTGTCCCTC 263  
QY 317 GTGSCCCAGCGCCCAACTCTCCGAGCTGTCGCGCTCTCCATCTACGTCACCCGGGC 376  
Db 264 GTCCGGAGGCGCCCGCCCGCGCGCGCTGCTGCTCTTCTGTCGTAACCCCGGC 323  
QY 377 CAATTCGCCCCCGAGGAGGACCTCTCCACCTACCCCTCCGACTTCGACGCGGAGTAAAG 436  
Db 324 CAGTTCCGCCCTCCGAGGACCTCTCGACGTAACCATCTGATTTGAGGCGGACCTGGGC 383  
QY 437 AAATCGCTGCTGCTTCCGCGCGGCTGACGTCGTTTTTCCATCCCGTAACTTTGACGAT 496  
Db 384 AAGCTCAGGCGCGCTCCCGCGCGGCTGAGCGTCTTCTGTCGCCAGAACTCTTACGAC 443  
QY 497 TACGGGAAGAACGCTGCTGCTGAC-----GTGCGAGAGGCTGGTGAATGGTGTG 547  
Db 444 TAGCTCAACGCGAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503  
QY 548 TGCGTTGAGAGTGGTCCGCGGACGAAAGTTGGGTGAGGGTTGAGAGCTGGAATTTGGGG 607  
Db 504 TTGAGGAGAGGCGCATGGGCGACGAGGCGTGGTGGTGGAGGTTGGAGAGGCG 563  
QY 608 CTGTGTGGGAAGAGCGAGGCGGCTTTCTTCAGAGGGTGGCGACTGTGTGACGAAAGTTG 667  
Db 564 ATGTGCGGGAAGAGCAGGCGCGCTTCTTTCGAGGGGTGGCCACTGTGTGTCACCAAGCTG 623  
QY 668 TTTAATATTGTGAGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727  
Db 624 TTCAACATTGTGAGCGGATGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683  
QY 728 CTTATTTCAGAGGATGCTTCGAGATCTTGATTTTCCATAAAGTGTAGTGTGCTGAAATA 787  
Db 684 ATCAATTCGCGGTTGT---GAATCTTGACTTTCATTAAGTGTAGTGTCTGAGTCT 740  
QY 788 ACAAGTGAATAAGTGGCTGGCAATGAGTTACGTAATGTGCACTTTTCACTGAGAG 847  
Db 741 ATCGAGATCATGATGGCTTGGCTAAGCTCAGCAATGTGCTCTCTCTCCTCACTGAAGA 900

QY	848	AGGAAAAGGCAC	TATCAATAAAATAAATCATTTGTTAAAGAGCTAAATCAGACGACGAGAT	907
Db	801	AGGGAAAAGGCA	TGTGTCATTAAGCAGGTGTCATGTCAGAGAGCTAAATCTGTCGCGAAAAG	860
QY	908	GGTCAGGTGCAT	TGTGGAGAAGTTCACAAATTTGGTCATCCAAAGTGTTACTGATGCTGGT	967
Db	861	GGTCAAGTTRAC	TGTCGCAAAATCTTAAGGACTCCGTCATCCAGCAATACAGGAAGCTGGT	920
QY	968	GGAGGATCGATT	ATGCTGAGATTTGGATCAAAATAATTTGGAGAAAGTGGAAACAGATC	1027
Db	921	GGGAGAGATTG	ATTATGCTGAGATTTGATGATCAAGAGAGTTTAGAGGCAGTTGAAGAAATC	980
QY	1028	AAGAGTCCTGTC	GCTCTCTGTTGTCGTCATGGTTGCCAAGTTCAGGCTTATAGACAAC	1087
Db	981	AGGAGCCCCGTC	GCTCTCTGTCGCGCTGGTTGGGAAGGTCAGGCTGATTGATAAC	1040
QY	1088	ATGGAATCAACT	TGTGTCATGAATGTTTGATCTAACCTTCTGTCATCTCAACATGGG	1145
Db	1041	ATAGAAATCAAT	GTATGAAGAACACCTTTGTGCTGCTCATCTCATGAGGAGGG	1098

RESULT 3

```

US-10-033-269-3
; Sequence 3, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BSI446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1831
; TYPE: DNA
; ORGANISM: Zea mays
US-10-033-269-3

```

Query Match	29.6%;	Score 389.8;	DB 13;	Length 1831;
Best Local Similarity	64.7%;	Pred. No. 1.3e-106;		
Matches 618;	Conservative 0;	Mismatches 322;	Indels 15;	Gaps 2;
QY	159	TCCAC	TGTTTCCA	ATATTATCAATGCTCCAGGCCCAAGGGTGATCTCCGACAGG 218
Db	212	TTCC	ATTTTGACCA	TGCGGCGCGCGGTTCATGAGCCGGAGTGATCCGCGACAGG 271
QY	219	CCTCG	ATGCGGAGCTGTCTCGGCTCGATCGGGCC	CAGGGCAAGCTCATTTGGGGTGTTCC 278
Db	272	CGCG	ATGCGCGGTGTCTCGCGCGCGCGGAGGAA	GCGCGTGCCTTTAGTCC 331
QY	279	CCACC	TGGGCTTCCTCCAGCGGGCCACCTCTCGCTCGTGGCC-----CAGGCC 329	
Db	332	CCACG	ATGGGCTTCCTCCAGGAGGACACCTCTCGCTCGTCTCGCGCGCGTGGCGGCT 391	
QY	330	GCACA	CTCTCCGACGTGTGTCGGTCTCCATCTACGTCA	ACCGGGGCCAATTGCCCCCCA 389
Db	392	CCGCG	GCCCGCTGCGCGCTCGTCTCTCATCTACGTCA	ACCCGAGCTGGCCCCCA 451
QY	390	CGAGG	AGCCTCTCCACCTACCCCTCGACATTCGACGGCGGACGTAA	GAAACTCGCTCCG 449
Db	452	CCGAG	GAACCTTGGCAACCTACCCCTCGACATTCGCGCGGACCT	CGGCAAGCTGGCGGCCA 511
QY	450	TTCC	CGCGCGCTGCGACGTGTTTCCATCCCGTAACTTGTACGATT	ACGGGAAGAAGC 509
Db	512	CCGGG	TGCTCGGCGAGTATTTTGCCCCCGGACCTCTACGTC	CGCGGAGCGCTGACC 571
QY	510	GTGGT	GTCAGCTGGGACAGGCTGTTGGAA	TGTTGTCGTGTTTAG-----AGTGGT 563

Db	572	GC	CCCTCGCGCGCCAGCGCTCCGGGGCGCGTGTCTTGCTGAGAGATCGAGTGGGC	631
Qy	564	CG	GGCAGCAAAATTTGGTGTAGGGTTTGAGAACTGGAAATTTGGGGCTGTGTGGGAAGACA	623
Db	632	AT	GGCACGAGACGTGGATTCCGGGTGGAGCGGCTGGAGAGGGGCTGTGCGGGACGACA	691
Qy	624	GG	CCCGCTTTCTTCAGAGGGGTGGCGACTGTGTGTGACGAACTGTGTTAATATTGTGGAGC	683
Db	692	GG	CCCTGTCTCTTCGCGGGCGTGGCCACCGTGTGCGCAAGCTGTTCAATATCTGTGGAGC	757
Qy	684	CAG	ATGTGGCTGTGTTCGTTAAGAGGATATACAGCAGTGTGCGGCTATTTCAGAGGATGG	743
Db	752	CG	GACGTGCGCGTGTTCGGGAAGAGGATTAACAGCAGTGGCGCGTCAATTTGTCGGATGG	811
Qy	744	TT	CAGATCTTGATTTTTCCATTAAGTGATAGTGTCTGTAATTAACAGTGATATGATG	803
Db	812	TT	CGTGTCTTTGATTTTGCCCATACAGATAGTGTGTCGGAGGTAGTGCAGAAAGCTGATG	871
Qy	804	GC	TGGCAATGAGTTCACGTAATGTGCACCTTCACCTGAAGAGAGGAAAAAGGCATAT	863
Db	872	GT	CTTGAATGAGCTCTCGCAATGTAATCTGTACAGAGGATAGAAAGAGCGGTAT	931
Qy	864	CA	ATAATAATCAATTGTAAGACTAAATCAGCAGCAGGAGATGGTCAGTGCATTGTG	923
Db	932	CG	ATCAGTAGATCACTGGTGGATGCTAGAACCGGCAACCTCAGTGTGAAGCAACCGTAGCC	991
Qy	924	AGA	AGTTGACAAATTTGGTCAATCAAAAGTGTTACTGATGCTGGTGGGAAGGATCGATTATG	983
Db	992	AAG	AGATAAAGATCAAAATAGTGGGACAAATTCAGAAGCTGGCGGTACAGTTGACTATG	1051
Qy	984	CT	GAGATTGTGATCAAAATAAATTTGAGAAAGTGGAAACAGATCAAGAGTCCCTCGCTCT	1043
Db	1052	TT	GAGATTGTGGAGCAGAAAGCTTGGTGCTGTGGAGAGATGGACCGCCCTGTGTGTA	1111
Qy	1044	TC	TGTTGTCGATGTTTTGGCAAGTCAAGCTCAGGCTTATAGACCAATCGGAATCAA	1098
Db	1112	TT	TGTGTCGGCGATGTTTTGGAAGGTCAAGGTAATTTGACCAATCGAGATCA	1166

```

RESULT 4
US-10-033-269-12
; Sequence 12, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 1235
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-033-269-12

```

	Query Match	29.2%;	Score 384.4;	DB 13;	Length 1235;
	Best Local Similarity	66.7%;	Pred. No. 4.1e-105;		
	Matches 620;	Conservative 0;	Mismatches 286;	Indels 24;	Gaps 4;
QY	187	GGTCTCAGCCCAAGGGTGATCTCCACAAGGCTCGATCCGAGAGTGTGCGCGCTCGAT	246		
Db	75	GGCGGGCGAGCCGGAGGTGATCCGGACAAAGCGCGGATCGGGCGTGTGCGGGCGGCCA	134		
QY	247	GGCGGGCCCAAGGCAAGCTCATTTGGGTGTGTCTCCCAACATGGGCTTCTCCACGCGGGCCA	306		
Db	135	GGCGGGCGAGGGAAGACCGGTGTGTCGTGTGCCCAACATGGGCTTCTCCACGAGGGGCCA	194		

```
QY 307 CTTCTCGCTCGT-----GGCCAGGCGCCCAACTCTCCGAGTGTGTCGCGCTCTCCAT 360
Db 195 CTTCTCGCTCGTCTCGCGCGCGCGCGCTGCGCGCGCGCTGCGCTGCTCTCCAT 254
QY 361 CTACGTCACCCCGGCGCAATTCGCCCGCAGGAGGACTCTCCACCTACCCCTCCGACTT 420
Db 255 CTACGTCACCCCGGCGCAATTCGCCCGCAGGAGGACTCTCCACCTACCCCTCCGACTT 420
QY 421 CGAGCGCGAGTAAAGAAATCCGCTCGCTCGCTTCCGCGCGGCTCGAGCTGTTTCCATCC 480
Db 315 CGCGGGGACTCCGCAAGCTCGCTCC-----ACGGGCGGCTCCAGCGCTCTTCAACC 371
QY 481 CCGTAATTGTACGATTACGGGAAGAACGCTGCTGTGACGCTGGCAGAGCT-----532
Db 372 CCAGACCTCTACCAACCGCGCGCGCTGCTCTGCGCGCGCGCGCGAGGCTCCCGCGG 431
QY 533 ----GGTGGNAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
Db 432 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
QY 589 TGAGAAAGCTGGAATTTGGGCTGTGTGGGAAGAGCAGCGCGCTTTCTTCTCAGAGGGGTGC 648
Db 492 GAGCGGCTGGGAAGGCTCTGTGGGCGCAGCGCGCGCTGCTTCTTCTGCGGTGGG 551
QY 649 GACTGTGGTGAAGAGTGTGTTAATATGTGGAGCCAGATGTGGCTGTGTGCGTAAGAA 708
Db 552 CACCGTGTGCGCAAGCTGTTCAACGCTGTGAGCGCGCGCGCGCGCTGCTGCGGAAGAA 611
QY 709 GGAATATCAGAGTGGCGCTTATTCAGAGGATGCTTCAGATCTTGTATTTTCCATAAA 768
Db 612 GGAATACAGAGTGGCGCTTATTCAGAGGATGCTTCAGATCTTGTATTTTCCATAAA 768
QY 769 AGTGATGCTGTGGAATAACAGTGAATATGATGCGCTGGCAATCAGTTCAGTAATGT 828
Db 672 GATATAGAGGAGAAATAGTGGAGAGCAGATGTCTTGCAATGAGCTCTCGCAAGT 731
QY 829 GCACCTTTCACCTGAAGAGGAAAGGCACTTATCAATATAATAATCAATTTTAAGAGC 888
Db 732 CCACCTCTCGCTGAGGAAGGAGAGGCACTTATCAATAGTAGTCACTGTTAAATGC 791
QY 889 TAAATCAGAGC---AGGAGATGCTCAGTGCATTTGTGAGAGTTCACAAATTTGTGCTAT 945
Db 792 TAGAATGCTGCTGCTGAATATAGCAACAGTGTAGCAACATATAAGGATCAGATAGT 851
QY 946 CCAAGTGTACTGATGCTGGTGGAGAGGATCGAATATGCTGAGATTTGTTGATCAATAA 1005
Db 852 GCAGACGCTGACTGAAGCTGGCGTGGGTTGATATGTTGAGATTTGTGAGCAGGAAG 911
QY 1006 TTTGGAGAAAGTGGAAACAGATCAAGAGTCTGTCGCTTCTGTTGTGCTGCTGTTGG 1065
Db 912 TTTGGTACCTGTGGAGACATCGACCGCGCTGTTGTCTTGTGCTGCTGCTGCTGTTGG 971
QY 1066 CAAAGTCAGGCTTATAGCAACATGCAAT 1095
Db 972 AAAGGTTAGATTGATGATAATATGCAAT 1001
```

## RESULT 5

```
US-10-033-269-1
; Sequence 1, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; PRIORITY FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
```

```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-033-269-1
```

```
Query Match 29.1%; Score 382.6; DB 13; Length 1443;
Best Local Similarity 65.0%; Pred No. 1.6e-104;
Matches 603; Conservative 0; Mismatches 309; Indels 15; Gaps 2;
```

```
QY 187 GCCTCAGCCCCAAGGTTATCTCCGACAAAGCCTTCGATGCGGAGCTGTGTCGCGCTCGAT 246
Db 144 GGTTCATGAGCGGAGGTTGATCCGCGACAAAGCGGCGATGCGCGTGTGTCGCGCGCG 203
QY 247 GCGGCGCCAGGCGAAGCTCATTTGGCTGTGTCCTCCATGCGCTTCTCCACGCGGGCA 306
Db 204 CGCGCGGAGGAAAGGCGCTGCTTTAGTCCCAAGATGGGCTTCTCTCCAGAGGACA 263
QY 307 CTTCTCGCTGCTGCGGCCAGGCGCGCAACTCTCG-----ACGTCGTGCGCTCTC 357
Db 264 CTTCTCGCTCATCTCCGCGCGCTGCGCGCGCTTCCCGCGCGCGCTGCTGCTCTC 323
QY 358 CATCTACGTTCAACCGCGGCCAATTCGCCCGCCACGAGGACCTCTCCACTACCCCTCCGA 417
Db 324 CATCTACGTTCAACCGCGGCCAATTCGCCCGCCACGAGGACCTCTCCACTACCCCTCCGA 383
QY 418 CTTGACCGCGCAGCTAAAGAAACTCGCTCGCTTCCCGCGCGCGCTGCTGCGCTGCTTCCA 477
Db 384 CTTGCGCGCGCAGCTTCCGCAAGCTGCGCGCGCGCTGCTGCGCGCTGCTGCTGCTTGTCC 443
QY 478 TCCCGCTAACTGTACGATTACGGAAGAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
Db 444 CCGACACCTCTACGTTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503
QY 538 AATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
Db 504 CGCGGTGCTTGTGCTGAGGATGCGAGTGGGATCGCGCACGAGAGCTGCTGCTGCTGCTG 563
QY 592 GAGCTGGAATTTGGGCTGTGGAAGAGCAGCGCGCTTCTTTCAGAGGCTGCGGAC 651
Db 564 CGCGCTGAGAGGAGGCTGTGCGGAGCAGCAGGCTGCTTCTTCTGCGCGCTGCTGCTG 623
QY 652 TGTGTGACGAGTGTGTTAATATTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
Db 624 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683
QY 712 TTTACGAGTGTGCGGCTTATTCAGAGGATGCTTCTGAGATCTTGTATTTTCCATAAAGT 771
Db 684 TTTACGAGTGTGCGGCTTATTCAGAGGATGCTTCTGAGATCTTGTATTTTCCATAAAGT 743
QY 772 GATAGTGTGCTGAATAACAGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db 744 AGTTGGTTCGGAAGTAGTTCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803
QY 832 CTTTTCACCTGAAGAGGAGGAAAGGCTATCAATAATAATAATAATAATAATAATAATA 891
Db 804 TCTGTACAGAGGAGATAGAAAGAGGCTTATCAATGATGATGCTGCTGCTGCTGCTGCTG 863
QY 892 ATCAGCAGCAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
Db 864 AACCGCGCGCTCAGTGAAGCAACCGTAGCAAGAAATAATAATAATAATAATAATAATA 923
QY 952 TGTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011
Db 924 AATTACAGAGCTGCGGCTGAGGTTGATGATGATGATGATGATGATGATGATGATGAT 983
QY 1012 GAAAGTGAACAGATCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
Db 984 CTTGTGAGAGGATGAGCCCGCTTGTGTAATTTGTGCTGCTGCTGCTGCTGCTGCTG 1043
QY 1072 CAGGCTTATAGCAACATGGAATCAA 1098
|||||
```

```
Db 1044 CAGGCTAATTGACACATCGAGATCCA 1070

RESULT 6
US-10-033-269-10
; Sequence 10, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 1148
; TYPE: DNA
; ORGANISM: Tulipa fosteriana
US-10-033-269-10

Query Match 26.3%; Score 346.2; DB 13; Length 1148;
Best Local Similarity 63.7%; Pred. No. 1.4e-93;
Matches 583; Conservative 0; Mismatches 263; Indels 69; Gaps 1;

QY 192 CAGCCCCAAGGTGATCTCCGACAAAGGCTCGATCGGAGCTGTGTCGGCTCGATGCGGG 251
Db 79 CGGAACCAAGTCATAATTAACCTCGAAACCCGAGATCTAGCTGTGTCGGGACACACCGGC 138
QY 252 CCCAGGGCAAGTCTAATGGGTGTGTCCCAACCATGGGTCTTCTCCACGCGGGCCACCTCT 311
Db 139 GCCTCTCCACACCATCGCCCTGTGTCCTCCACCATGGGTCTTCTCCACGCGGGCCATCTCT 198
QY 312 GCCTGTGGCCAGCGCGCCCACTCTCTCGAGCTGTGTCGGCTGTCTCCATCTCCATCTACGTC 371
Db 199 CCCTCATCTCCATCGCCCTCCCTCGCGGACCTCACCGTGTCTCCATCTACCTCAACC 258
QY 372 CGGGCCAATTCGCCCCACGAGGAGCTCTCCACTACCTCCCTCCGACTTCGACGGCGAG 431
Db 259 CCACCAAGTTCGCCCCCTCCGAGGAGCTTCGCCACCTACCCCGGACCTCGCGCGGAC 318
QY 432 TAAAGAAACTCGGTCCTGTTCCCGCGGCGTCGAGCTGCTTTTCATCCCGTAACTTGT 491
Db 319 TCGCAACCTCGGCGCTGCCCCCTCGTCGCGCGCTTCTTGCCCCACTAATCCCT--- 375
QY 492 ACGATTACGGAAGAACGGTGTGTGACGTGGCGAGGGCTGGTGGAAATGTTGTGTCG 551
Db 376 ----- 375
QY 552 TTGAGAGTGGTTCGGGCAAGAAAGTTTGGGTGAGGTTTGAGAGCTGGAATGGGCTGT 611
Db 376 -----ACCGGAACGGGCAAGAGAGTGGGTGAGGTTGAGAGCTGAGAGCGGGGCTGT 429
QY 612 GTGGGAAGCAGAGCGCCCTTTTCTTCAGAGGGGTGGCGACTGTGTGTGACGAAGTTGTTTA 671
Db 430 CGGGCTGAGCGCGCGGTGTTTTTTCGGGTGTTCGCAACGGTGTGTGGAAGCTGTTTC 489
QY 672 ATATTGTGGACCGAGATGTGGCTGTGTTTCGGTAAAGAGGATTAATCAGCAGTGGCGGCTTA 731
Db 490 ATTTGGTGGAGCCCGATCGCGGTGTTTCGGGAAGAGGATTTTCAGCAGTGGCGGTGA 549
QY 732 TTCAGAGATGGTTCGAGATCTGATTTTTCATTAAGATGATAGTGTCTGAAATAACAC 791
Db 550 TCGAAGAAATGGTACGCGATCTGATTTTCTGTAAGGATTTGTTGGATCTCAAAATAGTAC 609
QY 792 GTGATAATGATGGCTGCAATAGTTCACGTAATGTGCACCTTTCACCTGAAGAGGG 851
Db 610 GGGAGTTGATGACTCGCCATGAGCTCACGTAATGTTTCGCTTAACCTGAAGCGAG 669

RESULT 7
US-10-033-269-5
; Sequence 5, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-033-269-5

Query Match 22.8%; Score 300.2; DB 13; Length 1644;
Best Local Similarity 72.7%; Pred. No. 1.5e-79;
Matches 405; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY 197 CCAAGGTGATCTCCGCAAGGCTCTCGAGCTGTGTCGGAGCTGGTTCGGCGCGGCCAG 256
Db 144 CCGCTGTGTCATCACCGCAAGGCGGAGATCGGAGTGGTTCGAGGTCCGCGCGGCCAG 203
QY 257 GGCAAGCTCATTTGGGTGTGTCGCCACCATGGGCTTCTCCACGCGGGCCACCTTCGCTC 316
Db 204 GGCAGGACCATCGCCCTCGTGTGCCACCATGGGTTCCTCCACGAGGCGCACCTTCGCTC 263
QY 317 GTGCCCCAGAGCGCCGCAACTCTCGAGCTGTGTCGGCTCTCCATCTACGTCAACCCGCGC 376
Db 264 GTCCGGAGGCGCGCCCGCGCGCGCGCGCGCGCTGTGTCTCCGCTACGTGAACCCCGCGC 323
QY 377 CAATTGCCCCCAACGAGGAGCTCTCCACCTACCCCTCCGACTTCGAGCGGCGAGCTAAAG 436
Db 324 CAGTTCGCCCTCCGAGGAGCTCTCGACGTACCCATCTGATTCGAGGGCGAGCTGGGC 383
QY 437 AACTCGCGTCTGTTCCCGCGGCGTGCAGCTGTGTTTCCATCCCGCTAACTGTACGAT 496
Db 384 AAGCTCAGGGCGGTTCGCCGCGCGTGGACGTCTCTTCGTCGCCAGAACTCTTACGAC 443
QY 497 TACGGGAAGAACGGTGTGTGTGAC-----GTGCGAGAGCTGTGGAATGGTGTG 547
Db 444 TACGCTCAACGCGAGGTGCGGTGGCTCTGGGCTTGAGAGCGAATAATGGGTCCGTCTGT 503
QY 548 TGGCTTGAGAGTGGGTCCGGGCGACGAAAGTTGGGTGAGGCTTGAGAGCTGGAATTCGGG 607
```



Db 504 TTGAGAGAGAGAGGCGATGGGCGACGAGGCGTGGGTGAGGGTGGAGAGGTTGGAGAGGGC 563  
Qy 608 CTGTGTGGGAGAGACAGCCCGTTTCTTCAGAGGGTGGCGACTGTGGTGCAGCAAGTTG 667  
Db 564 ATGTGCGGGAAGAGACAGCCCGTGTCTTTCGAGGGGTGGCCACTGTGGTCAACAAGCTG 623  
Qy 668 TTTAAATATTGGAGCCAGATGTGGCTGTGTTCGGTAAGAAGATTATCAGCAGTGGCGG 727  
Db 624 TTCACATTTGGAGCCGGATGTGAGTGTTCGGGAAGAGGATTACCAAGCAGTGGCGG 683  
Qy 728 CTTATTTCAGAGGATGGT 744  
Db 684 ATCATTCGGCGGTGGT 700

## RESULT 8

US-10-114-170-7/c  
; Sequence 7, Application US/10114170  
; Publication No. US20030023075A1  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Valerie  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod  
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 8.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/114,170  
; FILING DATE: 01-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/453,702  
; FILING DATE: 03-DEC-1999  
; APPLICATION NUMBER: 60/110,955  
; FILING DATE: 04-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seay, Nicholas J.  
; REGISTRATION NUMBER: 27386  
; REFERENCE/DOCKET NUMBER: 960296.95017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000  
; TELEFAX: (608) 251-9166  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9381  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-114-170-7  
Query Match 9.0%; Score 118.8; DB 15; Length 9381;  
Best Local Similarity 48.3%; Pred. No. 4.1e-24;  
Matches 427; Conservative 0; Mismatches 397; Indels 60; Gaps 1;  
Qy 239 CGTCTGATGGGCGCCAGGCGCAAGCTCATTTGGCTGCCACCATGGGCTTCCTCCAC 298  
Db 1111 CGCGCGCTGTATGGAGGCAAGCGCGTGGCTGTACTTACCATGGGCAACCTGCAT 1052

Qy 299 GCAGGCGCACCTCTCGCTCGTGGCCAGGCGCCGCAACTCTCCAGCTCGTCGCGCTCTCC 358  
Db 1051 GATGGCCACATGAAGCTGGTCGACGAGCCAAAGCCGCGCGATGTGGTCTGCTCAGT 992  
Qy 359 ATCTAGTCAACCCGGGCGCAATTCGCGCCCAACGAGGACCTCTCCACTACCCCTCCGAC 418  
Db 991 ATTTCGTAAACCCGATGAGTTCGACCGCCCGGAGGACCTGCTCGGTACCCACGACC 932  
Qy 419 TTGACGGCGACGTAAGAACTCGCGTCCGTTCCCGGGCGTCCAGCTCGCTTTTCCAT 478  
Db 931 TTGACGAGACGCGCAAAAGCT----- 909  
Qy 479 CCCCCTAACTTGTACGATTACGGGAAGAACGGTGTGTGACGCTGCAGAGGCTGTGGA 538  
Db 908 -----GACAAAGCGCAAGTGTGATCTGTTTTTCGCGCCCTCG 872  
Qy 539 ATGGTGTGTGCGCTTGAAGTGGTCCGGGCGACGAAAGTTGGGTGAGGGTTGAGAAGCTG 598  
Db 871 GTAAAGAGATCTACCCGAAACGGTACTGAAACCATACTTACGTTGACGTTCCAGGGCTT 812  
Qy 599 GAATTGGGCTGTGTGGGAAGAGAGCGCGCTTTCTTCAGAGGGGTGGCGACTGTGGTG 658  
Db 811 TCGACCATGCTGAAGCGCCAGCCGCGGGGCAATTTTCGGGGCGTGTGCAATATCGTC 752  
Qy 659 ACGAAGTTGTTTAATTTGTGAGCCAGATGTGGCTGTGTTTCGTTAAGAGGATTAATCAG 718  
Db 751 AGCAAGCTGTTTAACTGTGTCAGCGGACATCGCTGCTTCGTTGAAAAGGACTTTTCAG 692  
Qy 719 CAGTGGCGGCTTATTCAGAGGATGTTTCGAGATCTTTGATTTTTCATAAAGTGTAGGT 778  
Db 691 CAACTGGCGCTGATCCGCAAAATGTTGCTGATATGGGCTTTGATATTGAGATTGTCGT 632  
Qy 779 GCTGAATAACACGCTGATTAATGATGCTGCGCAATGAGTTTTCAGTTATGTCACCTTCA 838  
Db 631 GTGCCAATTATGCGCGCCAAAGACGGTCTGGCACTGAGTTCGCGTAAACGTTATCTGACG 572  
Qy 839 CTGAAGAGAGGGAAGGCACTATCAATAAATAATCATTTGTTAAGAGCTAAATCAGCA 898  
Db 571 GCAGAACACGCAAAATTCGCGCGGCTGTGACAAAGTTTAAAGTTGATTCGATTCGCCGCAAA 512  
Qy 899 GCAGGATGTGTCAGGTGCAATTTGAGAAAGTTGACAAATTTGGTTCATCCAAAGTGTACT 958  
Db 511 TTGCGAGGCTGGCGAACGGGATCTCGATGAAATTAATTCCTATTTCGCGGGCAAGACTGAAT 452  
Qy 959 GATGCTGGTGGAGGATGATTAATGCTGAGATGTTGATCAAAATTAATTTGAGAAAGTG 1018  
Db 451 GAAAAGGCTTCGCGCGCGATGATTAATTCAGATTCGCGATGCCGACACATTTGCTGGAAGTC 392  
Qy 1019 GAACAGATCAAGAGTCTCTGCTCTTCTGTGTTGCTGATGTTTGGCAAGTCAGGCTT 1078  
Db 391 TCTGAAACCAGCAACGGGCGAGTAATTCGTGAGCGGCTGGCTGGCGATGCTCGCCCTG 332  
Qy 1079 ATAGACAAATGGAATCAACTTGTCAATGAATGTTTGTATCTAA 1122  
Db 331 ATCGACACAAAAATGTCGAGCTGGCGTAATACTTATATACTCTAA 288

## RESULT 9

US-09-070-927A-757/c  
; Sequence 757, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA



```

;
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <UNKNOWN>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 757:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1672 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 757:
US-09-070-927A-757

```

```

Query Match      8.3%; Score 109; DB 10; Length 1672;
Best Local Similarity 51.1%; Pred. No. 1.1e-21;
Matches 275; Conservative 2; Mismatches 258; Indels 3; Gaps 1;

QY 605 GGGCTGTTGGGAAGAGCAGCCCGCTTTCTTCAGAGGGGTGGCGACTGTGGTGAAGAAG 664
DB |||||
QY 593 GGTATTATGAGTAGTGGAGCCATACATTTTGTAGAGGGTTCACACTGTGTGAAGTAAR 534
DB |||||
QY 665 TTGTTATATATGTTGGAGCCAGATGTGGCTGTTCGTAAGAGGATTATCAGCAGTGG 724
DB |||||
QY 533 TTATTTAACTAATACACAGCAGATAGAGCTTATTTGGAGAAAAGGATGCACACAGTTA 474
DB |||||
QY 725 CGGCTATTTCAGAGATGGTTCGAGATCTTCATTTTCCATAAAAGTGATAGGTGCTGAA 784
DB |||||
QY 473 GCTGTAATTAAGAAATGTTAGGACCTTAATTTGATATAGACGTTGTAGGGTGCCA 414
DB |||||
QY 785 ATAACAGTGATATGATGGCTGGCAATGATGTTACGTAATGTCACCTTTTCACCTGAA 844
DB |||||
QY 413 ATTATAGGGAAGAGGATGGAGTACTAGCTAAAAGTTCTAGAAAACACATACCTTAAGTTAGAG 354
DB |||||
QY 845 GAGAGGGAAGAGGACTATCATATAATAATCATTTTAAAGCTAATTAAGCTAAGTAAAGTAAAG 904
DB |||||
QY 353 GAAGATCTTCAGCTACTATTTTAAATAAATTCYTTAATCTTAGCTAGGAGCCCTTAAT 284
DB |||||
QY 905 GATGTCAGGTGCAATGTGAGAAGTTGCAAAATTTGGTCAATCCAAAGTGT---TACTGAT 961
DB |||||
QY 293 AATGGAGAGAGATAGCTYAAAATAATKAGATAATAAGTAAAATAATAATAATTAATTAAT 234
DB |||||
QY 962 GCTGTTGAAGATCGATATGCTGAGATTTGTTGATCAAAATATTTGGAGAAAGTGA 1021
DB |||||
QY 233 AATTTGGCAAGATTTGATTTATGAGAGTTGTTGATCTTTATCATTAACAAGAGTAAT 174
DB |||||
QY 1022 CAGATCAAGAGTCCGTCTCTTCGTGCTGCTGTTGCAAGTGGTGGCAAGTCAAGCTTATA 1081
DB |||||
QY 173 TATATTGAATAATCAGTTTTAGTTCCAATAGCAGTATTTATAGAAAATAACAGATTATA 114
DB |||||
QY 1082 GACACATGGAAATCAACTGTCAATGAATGTTGATCTCAACCTTCTGTCACTCAAA 1139
DB |||||
QY 113 GATAATTTTACATTTGAATTAACAATAGGGGGCTATAAATGANAAGTTCATGTTAA 56
DB |||||

```

## RESULT 10

```

US-08-781-986A-174/C
; Sequence 174, Application US/08781986A
; Publication No. US2003005436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4549 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-174

```

Query Match 7.5%; Score 99; DB 8; Length 4549;

Best Local Similarity 51.5%; Pred. No. 2.4e-18;

Matches 250; Conservative 1; Mismatches 231; Indels 3; Gaps 1;

```

QY 614 GGGAGAGCAGCCCGCTTTCTTCAGAGGGGTGGCGACTGTGGTCAAGAGTATATCAGCAGTGGCGCTTATT 673
DB |||||
QY 3351 GGAGCGAAGCGYCCAGGCGCATTTTGATGGGGTGGTAAACAGTAGTTAATAAGCTATTTTAAT 3292
DB |||||
QY 674 ATTGTGGAGCCAGATGTGGCTGTGTTTCGTAAGAGGATTATCAGCAGTGGCGCTTATT 733
DB |||||
QY 3291 ATTGTTATGCCGATTATGCTTATTTTGTGTAAGAGATGCCAGCAATTGGCTATTGTC 3232
DB |||||
QY 734 CAGAGATGTTTCGAGATCTTGATTTTCCATAAAGTGAAGTGTGCTGTAATTAACAGCT 793
DB |||||
QY 3231 GAGCAAAAGGTAAAAGACTTCAATCATCCCGTTGAAATTTATTTGTTATGATATCGTTGGA 3172
DB |||||
QY 794 GATAATGATGGCTCGCAATGAGTTCAGTATGTGACCTTTTCACCTGAAGAGAGGAA 853
DB |||||
QY 3171 GAAGCAGATGGTTGGCGAAAGTTCAAGAAATGTTTATTAACGGACCAAGACGACAA 3112
DB |||||
QY 854 AAGGCACATCAATAAATAAATCAATTTGTAAGAGCTAATAATCAGCAGCAGGAGATGGTCAG 913
DB |||||
QY 3111 GAAGCGGTACATTTAAGTAAAGTTTGTATTAGCACAAGCGTTGTATCAAGACGGTGAA 3052
DB |||||
QY 914 GTGCATTGTGAGAAGTTGCACAAATTTGCTCATCCAAAGTGT---ACTGATGCTGGTGA 970
DB |||||
QY 3051 CGTCAAGTAAAGTAATTAATATAGCAGGGTTACTGAAATATCTTGAATCATCATATAGTGAA 2992
DB |||||
QY 971 AGGATCGATTATGCTGAGATTTGTTGATCAAAATAATTTGGAGAAAGTGGACAGATCAAG 1030
DB |||||

```

Db 2991 CCATTGAGAGAGTTGCTGTTTATAGTTATCTCAATTAGTAGAACAACAGAAATAACT 2932  
QY 1031 AGTCCTGTCGCTTCTGTTGTTGTCATGTTGGCAAGTCAGGCTTATACACACATG 1090  
Db 2931 GGACGCATATTATCTCGTTAGCTGTAAATTTTCAAGAGCGGTTTATAGATATATA 2872  
QY 1091 GAAAT 1095  
Db 2871 ATAAT 2867  
RESULT 11  
US-09-974-300-1035  
; Sequence 1035, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1035  
; LENGTH: 546  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-1035  
Query Match 6.4%; Score 83.6; DB 10; Length 546;  
Best Local Similarity 47.8%; Pred. No. 2.5e-14;  
Matches 242; Conservative 0; Mismatches 264; Indels 0; Gaps 0;  
QY 584 AGGTTGAGAAAGTGGGAATGGGGCTGTGTGGGAAGACGAGCGCCGCTTTCTTCAGAGGG 643  
Db 27 AGCGTGAAAGCGGACGGATGTCTTTGCGCGCTTCAAGAGAGGGCATTTGACGGT 86  
QY 644 GTGGCGAGTGTGGTACGAATGTTTAATATTGTGGAGCGAGATGTGGCTGTGTCGGT 703  
Db 87 GTCGCAACAGTGTCTCAAAAGCTGTCAATCTCAGTCTCCGACAAAGSGTTTATTTCGGC 146  
QY 704 AAGAGGATTATACGACGTGCGGCTTATTTCAGAGGATGTTTCAGATCTTGATTTTCC 763  
Db 147 ATGAAGGATGCCAGCAGGTTGCGCTGCTGACGGGCTGATCAACGACTTTTTCATGGAT 205  
QY 764 ATAAAGTGATAGTGTGTAATAACACGTGATAATGATGCGCTGGCAATGAGTTCAGT 823  
Db 207 ATCAGCTGATCCCTGTGCAAAACAAACGGGAAGACGGCTTGCCTCAAGAGCTCAAGA 266  
QY 824 AATGTGACCTTTCACTGAAGAGAGGAAAGGCACTATCAATAAATAATCAATTGTTA 883  
Db 267 AATGCTAATCTGCGAGCTGAAGAACGGCAGGAGCAACCGCTCTTTACCGAGCTCTTCAG 326  
QY 884 AGAGCTAAATCAGCAGCAGGAGATGGTTCAGTGCATTTGTGAGAGTTGACAAATTTGTC 943  
Db 327 CGCGGCCGAGCTGATCCGCAATGGAGAGAGAGACCCGGAGGCGGTTAAGCGGAGATT 386  
QY 944 ATCCAAAGTGTACTGATGCTGTGGAGGATCGATTATGCTGAGATTGTTGATCAAAAT 1003  
Db 387 CGGAGCATCTGGAAAACAGTGGTGTATTGATTTATGCGGATATTTATTCATATCCC 446  
QY 1004 AATTGGAGAAAGTGGAAACAGATCAAGAGTCTGCTCTCTGTGTGCTGTCGATGTTT 1063  
Db 447 GATCTGAGATCAGGGATCCATTGACTGGAAGATCATCATAGCGGTTCGGTTCAGTTT 506  
QY 1064 GGCAAGCTCAGGCTTATAGCAACAT 1089  
|||||

Db 507 TCCAAAGCCCGCTTAATTGACAATAT 532  
RESULT 12  
US-09-790-988-1/c  
; Sequence 1, Application US/09790988  
; Patent No. US20020127687A1  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEMI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1  
Query Match 5.6%; Score 73.4; DB 10; Length 640681;  
Best Local Similarity 55.8%; Pred. No. 4.3e-09;  
Matches 140; Conservative 0; Mismatches 111; Indels 0; Gaps 0;  
QY 624 GGCCCGCTTTCTTCAGAGGGGTGGCGACTGTGGTGACGAAAGTTGTTTAAATATTGTGGAGC 683  
Db 213389 GACCTGGACACTTTAGAGGCGTTACGCAATTTACTAAGATTATTAACTTTATACAAC 213330  
QY 684 CAGATGGCTGTGTTTCGTTAAGAAGATTATCAGCAGTGGCGGCTTTATTCAGAGGATGG 743  
Db 213329 CAGATTTTGCATTTTGGAGAAAAAGATTATCAACAATTTATTAATAAATCCCTTG 213270  
QY 744 TTCGAGATCTTGATTTTCCATAAAGTGTGCTGCTGAAATAACACGCTGATAATGATG 803  
Db 213269 TAAAGGATTTGAATTATATGATAAAAAATAATTAGCTTACCTACATACGACTAAAAATG 213210  
QY 804 GCCTGGCAATGAGTTCACGTAATGTGCACCTTTCACTGAAGAGAGGAAAGGCACTAT 863  
Db 213209 GATTAGCTTTAAGTTCAGGAATAATTTAAGTTCCCAAGAAATGAAATAGCACCTT 213150  
QY 864 CAATAAATAAA 874  
Db 213149 ATTTATATAAA 213139  
RESULT 13  
US-09-738-626-132  
; Sequence 132, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 132  
 ; LENGTH: 837  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-132

Query Match 5.4%; Score 70.8; DB 10; Length 837;  
 Best Local Similarity 55.8%; Pred. No. 2.5e-10;  
 Matches 135; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 620 AGCAGGCGCGTTTCCTCAGAGGGGTGGCGACTGTGGTGAAGGATTTAATATTGG 679  
 Db 358 AGCAGGCGCGTTTCCTCAGAGGGGTGGCGACTGTGGTGAAGGATTTAATATTGG 679  
 QY 680 GAGCCAGATGGCTGTTCGATGTTGGTAAAGGATTTAATATTGG 739  
 Db 418 GAGCCAGATGGCTGTTCGATGTTGGTAAAGGATTTAATATTGG 739  
 QY 740 ATGCTTCGAGATCTTGTATTTTCCATAAAGTGTGACCTTCAACCTGATAAT 799  
 Db 478 ATGCTTCGAGATCTTGTATTTTCCATAAAGTGTGACCTTCAACCTGATAAT 799  
 QY 800 GATGCGCTGGCAATGAGTTCACCTTAATGTCACCTTCAACCTGATAAT 859  
 Db 538 GATGCGCTGGCAATGAGTTCACCTTAATGTCACCTTCAACCTGATAAT 859  
 QY 860 CT 861  
 Db 598 CT 599

## RESULT 14

US-09-738-626-1/c  
 ; Sequence 1, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; PRIOR FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3309400  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-1

Query Match 5.4%; Score 70.8; DB 10; Length 3309400;  
 Best Local Similarity 55.8%; Pred. No. 8.4e-08;  
 Matches 135; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 620 AGCAGGCGCGTTTCCTCAGAGGGGTGGCGACTGTGGTGAAGGATTTAATATTGG 679  
 Db 126832 AGCAGGCGCGTTTCCTCAGAGGGGTGGCGACTGTGGTGAAGGATTTAATATTGG 679  
 QY 680 GAGCCAGATGGCTGTTCGATGTTGGTAAAGGATTTAATATTGG 739  
 Db 126772 GAGCCAGATGGCTGTTCGATGTTGGTAAAGGATTTAATATTGG 739  
 QY 740 ATGCTTCGAGATCTTGTATTTTCCATAAAGTGTGACCTTCAACCTGATAAT 799  
 Db 126712 ATGCTTCGAGATCTTGTATTTTCCATAAAGTGTGACCTTCAACCTGATAAT 799  
 QY 800 GATGCGCTGGCAATGAGTTCACCTTAATGTCACCTTCAACCTGATAAT 859  
 Db 126652 GATGCGCTGGCAATGAGTTCACCTTAATGTCACCTTCAACCTGATAAT 859  
 QY 860 CT 861  
 Db 126592 CT 126591

## RESULT 15

US-09-974-300-5461  
 ; Sequence 5461, Application US/09974300  
 ; Patent No. US20020146721A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berka, Randy M.  
 ; APPLICANT: Clausen, Ib Groth  
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
 ; TITLE OF INVENTION: Expression  
 ; FILE REFERENCE: 10085.500-US  
 ; CURRENT APPLICATION NUMBER: US/09/974,300  
 ; CURRENT FILING DATE: 2001-10-05  
 ; PRIOR APPLICATION NUMBER: 09/680,598  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: 60/279,526  
 ; PRIOR FILING DATE: 2001-03-27  
 ; NUMBER OF SEQ ID NOS: 8481  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5461  
 ; LENGTH: 501  
 ; TYPE: DNA  
 ; ORGANISM: Bacillus clausii  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(501)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-974-300-5461

Query Match 5.1%; Score 67.2; DB 10; Length 501;  
 Best Local Similarity 53.9%; Pred. No. 2.2e-09;  
 Matches 138; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 605 GGGCTGTGTGGGAAGAGCAGCGCGCTTTCTTCAGAGGGGTGGCGACTGTGGTCAAG 664  
 Db 227 GTGTATCGCGAAGAGCGCGCTTGGCCATTTTCAGAGGGGTGGCGACTGTGGTCAAG 664  
 QY 665 TTGTTTAATATTGTGGAGCCAGATGTGGCTGTTCGGTAAAGAGGATTTATCAGAGTGG 724  
 Db 287 TTGTTTAATATTGTGGAGCCAGATGTGGCTGTTCGGTAAAGAGGATTTATCAGAGTGG 724  
 QY 725 CGGCTTATTCAGAGGATGGTTCGAGATCTTGATTTTCCATAAAGTGTAGTGTGAA 784  
 Db 347 GCCATTATTAAACAACTTGTTAAGGAGTTCCTTCCTGCTGTGACAAATTTGTGTTGCCA 406  
 QY 785 ATAACACGTGATTAATGATGGCTGGCAATGAGTTTCACTATGTGCACCTTTCACTGAA 844  
 Db 407 ACTGTTCCGAGAGCAGCGCTTGGCGAAGAGTCCCGCAACGCAATTTGACCGCGCA 466  
 QY 845 GAGAGGGAAGAGCGAC 860  
 Db 467 GAACGCCAAGCGCGAC 482

Search completed: December 16, 2003, 01:41:46  
Job time : 501 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 21:30:25 ; Search time 97 Seconds  
(without alignments)  
5988.250 Million cell updates/sec

Title: US-10-033-269-8  
Perfect score: 1316  
Sequence: 1 tgggcacgaggtttctcag.....tggcattgtattgtctagg 1316

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCITUS COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	118.8	9.0	9381	4	US-09-453-702B-7
C 2	108	8.2	885	4	US-09-107-532A-1499
C 3	99.8	7.6	867	4	US-09-134-001C-1484
C 4	93.8	7.1	852	4	US-09-328-352-3790
C 5	91.8	7.0	579	4	US-09-252-991A-10716
C 6	91.8	7.0	891	4	US-09-252-991A-10624
C 7	91.8	7.0	909	4	US-09-252-991A-10450
C 8	91.8	7.0	1422	4	US-09-252-991A-10234
C 9	79	6.0	3717	4	US-09-221-017B-1095
C 10	70.8	5.4	2164	3	US-09-318-794A-1
C 11	70.8	5.4	2164	3	US-09-318-793A-3
C 12	69.2	5.3	4403765	3	US-09-103-840A-2
C 13	69.2	5.3	4411529	3	US-09-103-840A-1
C 14	56.4	4.3	1926	4	US-09-249-585A-4
C 15	54.4	4.3	1931	2	US-09-130-114-2
C 16	54.4	4.1	7218	1	US-08-232-463-14
C 17	54.2	4.1	619	3	US-08-998-416-842
C 18	53	4.0	7218	1	US-08-232-463-14
C 19	51	3.9	390	3	US-09-197-649-7
C 20	49.2	3.7	468	4	US-09-252-991A-11520
C 21	49.2	3.7	1419	4	US-09-252-991A-11636
C 22	49.2	3.7	1587	4	US-09-252-991A-11720
C 23	49.2	3.7	2304	4	US-09-252-991A-11802
C 24	49	3.7	390	3	US-09-197-649-7
C 25	47.8	3.6	465	4	US-09-252-991A-12201
C 26	47.8	3.6	936	4	US-09-252-991A-12266
C 27	47.8	3.6	4884	4	US-09-252-991A-12126

C 28	47.8	3.6	4884	4	US-09-252-991A-12292	Sequence 12, Appl
C 29	47	3.6	1476	4	US-09-434-288-12	Sequence 58, Appl
C 30	46.8	3.6	1749	4	US-09-450-072-58	Sequence 58, Appl
C 31	46.8	3.6	1749	4	US-09-351-348-58	Sequence 58, Appl
C 32	46.4	3.5	4674	4	US-09-410-551B-26	Sequence 30, Appl
C 33	46.4	3.5	4737	4	US-09-410-551B-30	Sequence 30, Appl
C 34	46	3.5	1749	4	US-09-450-072-56	Sequence 56, Appl
C 35	46	3.5	1749	4	US-09-351-348-56	Sequence 56, Appl
C 36	45.6	3.5	1289	2	US-08-344-833-1	Sequence 1, Appl
C 37	45.6	3.5	4190	3	US-08-924-345-1	Sequence 1, Appl
C 38	45.4	3.4	1458	3	US-09-105-537-9	Sequence 9, Appl
C 39	45.4	3.4	1565	3	US-09-320-878-24	Sequence 24, Appl
C 40	45.4	3.4	1565	3	US-09-141-908-22	Sequence 24, Appl
C 41	45.4	3.4	1565	3	US-09-657-440-24	Sequence 24, Appl
C 42	45.4	3.4	13613	3	US-09-105-537-3	Sequence 3, Appl
C 43	45	3.4	933	3	US-09-105-390-43	Sequence 43, Appl
C 44	45	3.4	1008	3	US-09-105-390-59	Sequence 59, Appl
C 45	45	3.4	2810	3	US-09-105-390-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-09-453-702B-7/c  
; Sequence 7, Application US/09453702B  
; Patent No. 6365723  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Nicole T.  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod  
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 9381

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-453-702B-7  
Query Match 9.0%; Score 118.8; DB 4; Length 9381;  
Best Local Similarity 48.3%; Pred. No. 1.2e-22;

Matches 427; Conservative 0; Mismatches 397; Indels 60; Gaps 1;  
QY 239 CGCTCGATCGCGCCAGGCAAGCTCATTTGGCTGTGTCCTCCACCATGGCTTCTCCAC 298  
Db |||||  
1111 CGCCGCTTGGTATGGAAGCAAGCGGTGGCTGTGTACCTACCATGGCAACTGAT 1052  
QY 299 GGGGCGCACTTCTCGTGTGGCCAGGCGGCCCAACTCTCGAGCTGTGCGCTGCC 358  
Db |||||  
1051 GATGCCACATGAAGCTGTGACGAAGCAAGCGCGGATGTGTGCTGCTCAGT 992  
QY 359 ATCTACGTCAACCCCGGCCAATTCGCCGCCAGGAGACCTCTCCACCTACCCCTCCAC 418  
Db |||||  
991 ATTTTCGTTAACCCGATGAGTTGACCGCCCGGAGGACCTGCTCGGTACCCAGCAC 932  
QY 419 TTGCGAGCGGAGTAAAGAACTCGCTGCTCCGTTCCCGGGCGGTGCACTGCTTTCCAT 478  
Db |||||  
931 TTGCAGGAAGACTCGAAAAGCT----- 909  
QY 479 CCCGTAACTTGACGATTAACGGGAAGAACGGTGTGTGAGTGGAGCGGTGGTGA 538  
Db |||||  
908 -----GAACAAGCGCAAGTGGATCTGTTTTCGCGCCCTCG 872  
QY 539 ATGTTGTGCTGCTGAGATGGGTCCGGGCAAGAAAGTTGGGTGAGGTTGAGAGCTG 598  
Db |||||  
871 GTAAAGAGATCTACCCGACGGTACTGAACCCATCTTACGTTCAGCTTCCAGGGCTT 812  
QY 599 GAATYGGGGTGTGFGGAAGAGCGGCCGCTTTTCITTCAGAGGGGTGGCGACTGTGGT 658  
Db |||||  
811 TCGACCATGCTGGAAGCGCCAGCGCGCGGGCATTTTCGCGGCTGTGCACTATCTC 752  
QY 659 ACGAGTTGTTTAAATATTGTGAGCCAGATGTGGCTGTGTTGCGTGAAGAGATTCAG 718  
Db |||||  
751 AGCAAGCTGTTTAACTGCTCCAGCGGACATCGCTGCTTCCGTGAAAGGACTTTT 692  
QY 719 CAGTGGCGGCTTATTCAGAGATGTTTCAGATCTTGAATTTTCCATAAAGTGTAGGT 778  
Db |||||  
691 CAACCTGGCGCTGATCCGCAAAATGTTGCTGATATGGCTTTGATTTGAGATGTGCGT 632  
QY 779 GCTGAAATPAAACGCTGATATGATGCTGCAATGATTCACGTAATGTGCACTTTTCA 838  
Db |||||  
631 GTGCCAATATATGCGCGCCAAAGACGGTCTGCACTGAGTTCCTCGTAACTGCTGAG 572  
QY 839 CCTGAAGAGAGGAAAGGCACTATCAATAAATCAATTTGTTAAAGCTTAAATCAGCA 898  
Db |||||  
571 GCAGAACCAACCAAAATTTGCGCCGCTGTACAAAGTTTAAAGTTGATTCGCGACAAA 512  
QY 899 GCAGAGATGCTCAGGTGCTATTTGAGAGTTGACAAATTTGGTATCCCAAGTTTACT 958  
Db |||||  
511 TTGAGGCTGCGGAACGGATCTCGATGAATTTATGCTATTCGCGGCAAGACTGAAT 452  
QY 959 GATGCTGTGGAAGGATCGATTATCTGAGATTTGATCAAAATAATTTGGAGAAAGTG 1018  
Db |||||  
451 GAAAAAGGCTTCCGCGCGATGATATTCAGATTCGCGATCGGACACATTTCTGGAATC 392  
QY 1019 GAAAGATCAGAGTCTGCTGCTTCTGTTGCTGATGCTGCTGCTGCTGCTGCTGCTG 1078  
Db |||||  
391 TCTGAACCAACCAACCGGCGATTAATCTGCTAGCGGCTGCTGCTGCTGCTGCTGCTG 332  
QY 1079 ATAGACACATGGAATCAACTTGTCAATGATTTTGTATCTAA 1122  
Db |||||  
331 ATGACACAAATAGGTCGAGCTGCGGTAACTTACTTACTCTAA 288

RESULT 2

US-09-107-532A-1499  
; Sequence 1499, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1499:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 895 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...885  
SEQUENCE DESCRIPTION: SEQ ID NO: 1499:  
US-09-107-532A-1499

Query Match 8.2%; Score 108; DB 4; Length 885;  
Best Local Similarity 52.4%; Pred. No. 3.6e-20;  
Matches 262; Conservative 0; Mismatches 235; Indels 3; Gaps 1;  
QY 605 GGGCTGTGTCGGGAAGACAGCGCCCGTTTCTTCAGAGGGGTGGCGACTGTGGTGACGAAG 664  
Db |||||  
379 GGTTTATGTGGAGCTAGTAGGCCCAATACATTTTAGAGGGGTTTGCACCTGTTGTAAGTAA 438  
QY 665 TTGTTTAATATTGTGGAGCCAGATGTGGCTGTTCGTTAAGAGGANTATCAGCAGTGG 724  
Db |||||  
439 TTAATTTAACATAATACCCAGCAGATAGAGCTTAATTTGGAGAAAAGGATGCACACAGTTA 498  
QY 725 CGGCTTATTCAGAGGATGTTTCAGATCTTGAATTTTCCATAAAGTATAGGTGCTGAA 784  
Db |||||  
499 GCTGTAATTAAGAAGATGTTAGGACCTTAATTAATATAGACGTTGTAGGGGTGTCGA 558  
QY 785 ATAACACGTGATTAATGATGGCCCTGGCAATGAGTTTCAGTAATGTGCACCTTTTCCCTGAA 844  
Db |||||  
559 ATTATAAGGAAGAGGATGGACTAGCTAAAGTTCTAGAAACACATACCTTAAGTTTAGAG 618  
QY 845 GAGAGGGAAGGCACTATCAATAATAATCAATGTTAAGAGCTAAATCAGCAGCAGGA 904  
Db |||||  
619 GAAAGATCTTTCAGCTACTATTTTAAATAAATCTTTAACTTTAGTAAAGAGCTTTAAAT 678  
QY 905 GATGGTCAGTGCATTGTGAGAGTTGACAAATTTGCTCATCAAAGTGT---TACTGAT 961  
Db |||||  
679 AATGGAGAGAGACAGCTTTAAATAAATTAATGAGATAATAAGTAAATAATAATACTTAT 738  
QY 962 GCTGGTGGAGGATCGATTATGCTGAGATTTGTTGATCAAAATAATTTGGAGAAAGTGA 1021

Db	739	AAATTGGCGAAGATTGATTATGTAAGAATTGTTGATCTTTATCATTAACAAGAGTAAT	798
Qy	1022	CAGATCAAGAGTCTCTGCGTCTCTGTGTGCTGCATGCTTTGGCAAGTCAAGCTTATA	1081
Db	799	TATATTGGAAAATCAGTTTGTAGTTGCAATGACAGTATTTATAGGAAAAACAAGATTATA	859
Qy	1082	GACAAACATGGAAATCAACTT	1101
Db	859	GATAATTTTACATTTGAATT	878

### RESULT 3

```

US-09-134-001C-1484
; Sequence 1484, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1484
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1484

```

Query Match	7.6%	Score 99.8	DB 4	Length 867
Best Local Similarity	51.1%	Pred. No. 6.6e-18		
Matches 261	Conservative 0	Mismatches 247	Indels 3	Gaps 1
Qy	582	TCAGAGGTTGAGNAGCTGGAATTTGGGCGTGTGTCGGAGAGCAGGCCGCTTTTCTTCAGAG	641	
Db	329	TGAAGAGTTGGACACTTGTGGCACAGTATTAGAGGGAGCACAAAGACTCGACACTTCGAG	388	
Qy	642	GGGTGGCGACTCTGTGTACGAAGTGTGTTAATATTCTGGAGCCAGATGTGGCTGTGTCG	701	
Db	389	GTGTGTGACCGTGTCAACAAACTATTATTAATTGTGCAACAGATTTATGCTATTTTG	448	
Qy	702	GTAAGAAGATTATCAGCAGTGGCGGCTTATTCAGAGAGTGGTTCGAGATCTTGATTTTT	761	
Db	449	GGAAAAAGGATGCACAAACAATTAGCTATTCTTGAAAAGATGGTTAAAGACTTTAATCTTC	508	
Qy	762	CCATAAAGTGTATAGGTGCTGAAATAACAGCTGATTAATGATGCGCTGGCAATGAGTTCAC	821	
Db	509	CTGTACATGTTATCGGTATTGATTCGTAGAGAGAAAAGATGGTTTAGCCAAAGCTCTA	568	
Qy	822	GTAATGTGCACTTTCACTCGAAGAGAGGGAAGGCACCTATCAATAAATCAATCATGT	881	
Db	569	GAAATATTACTTTGTGACCTCTGAAGACGAAAGAGGCAAAACATTTATATCAAGCTCTAC	628	
Qy	882	TAAAGACTAAATCAGACACAGAGATGGTCAGTGCTGATTCGTGAGAAGTTGACAAATTTGG	941	
Db	629	GCTTAGCAAGAAATTTGTATGAAGCGGGTGAACAGATAGCAATGAGATTATAGGTCAAA	688	
Qy	942	TCATCCAAAGTCT---TACTCATCTCGTGGAGGATTCGATTATGCTGAGATGTGTGATC	998	
Db	689	TCGCTCGGATTTTAAACAAAAATTTAGTGGACATATTGATGATTTGGGTATTTATAGTT	748	
Qy	999	AAAAATAATTTGGAGAAAGTGGAAACAGATCAAGAGTCTGTGCTCTCTGTGTGCTGCAT	1058	
Db	749	ATCCAAATCTTATACAACAATCAAGATTCATGAGCAATTAATCATCATTTGCGAGTTA	808	
Qy	1059	GGTTTGGGAAAGTCAGGCTTATAGACAAAT	1089	
Db	809	AAATTTTCTAAAGCAAGATTGATAGATATAT	839	

## RESULT 4

```

RES0011 4
US-09-328-352-3790
; Sequence 3790, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3790
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3790

```

Query Match	7.1%	Score 93.8;	DB 4;	Length 852;
Best Local Similarity	52.1%;	Prod. NO. 3e-16;		
Matches 259;	Conservative 0;	Mismatches 232;	Indels 6;	Gaps 2;

  

QY	609	TGTTGGCAAGACAGCAGCCCGT	TTTCTTCAGAGGGGTGGCAGCTGCTG	TGTGACGAAGTTGT	668
DB	353	TATGTGGAAAGCTCAGTCCGGGCA	NTTTGATGGCGTGGCACTGGTTGT	TGTACCAAACTCT	412
QY	669	TTAATAATTGTGGAGCCAGATGTG	CGCTGTGTTTCGGTAAAGAAGATTAT	CAGCAGTGGCGGC	728
DB	413	TTAATAATTGTGCACCCGNAITAT	GTCTTCTTTGGTCAAAAAGACTACC	CAGCAATTAGCAG	472
QY	729	TTATTTCAGAGATGGTTTCGAGAT	CTTGATTTTTTCCATAAAAGTGATAG	TGCTGTAATAAT	788
DB	473	TGATTTCGCCAGTTTGTACAGGAT	TGCAATATCCGTTTGAAGTTAT	TGGCGTACCAATTG	532
QY	789	CACGTGATAATGATGGCTGGCAAT	AGTGTACAGTAAATGTGCACCTTTC	CACCTGAAGAGA	848
DB	533	TTCTGTGCAGAGACGGTTTGGCA	CTGAGTTTCAAGAAATGGTTACT	TGTCTCTGAAACAGC	592
QY	849	GGGAAAGAGGCCTATCAATRAAT	AAATCATTTGTTAAAGAGCTAAAT	CAGCAGCAGGAGATG	908
DB	593	GTACAGTTGCACCACTCATTTAT	CAAGTTTAAAAACAGCGAGAGCA	ATTTACACCAG	652
QY	909	GTCCAGTGCAATGTGAGAGTTGTG	CAAAATTTGCTCATCCAAAAGTGT	TACTGATGCTGTG	968
DB	653	GTAAGAATTTACAACAGTCTTGG	CAGATTG---AAAACACTGTTG	CAGACAAATGGT	709
QY	969	GAAGGATCGAATGCTCAGATGTT	GATCAAAATTAATTTGGAGAAAG	TGGNACAGATCA	1028
DB	710	TTGTTGTGGATTATGTGGAG---	GCTCGTCAACCTAATTTGTGTAG	CTGTTCTCAATTG	766
QY	1029	AGAGTCCTGTGCTCTTCTGTG	TTCGTCTGGCAAGTCAGGCTTAT	TAGACAACA	1088
DB	767	ATCTGATATTGTGTGTTGTTG	TTCGGCGCAAAATTAGGTGGA	ACGCGCTTAATTGATATC	826
QY	1089	TGGAATCAACTTGTCA			1105
DB	827	TACAAGTTGCCTTTTACA			843

## RESULT 5

```

RESOLUTION 5
US-09-252-991A-10716/c
; Sequence 10716, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ARUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

```





QY 308 CTCTGCTGTCGCCCGCCGAGCCGCGCCAACTCTCCGAGCTGCTGCCGCTCTCCATCTACGTC 367  
Db 682 GCCGCGTGTGAAGAAAGCGCGGAAACGCGCGACTTCGTGCTCGTCAGCATCTTCGTC 741  
QY 368 AACCCGGGGCAATTCCGCCCGCCGAGGACCTCTCCACCTACCCCTCCGACTTCGACGGC 427  
Db 742 AACCGCTGCAATTCGCCCGCCGAGGACCTCGACAATACCCGGGACCCCTCGAGCC 801  
QY 428 GAC 430  
Db 802 GAC 804

RESULT 9  
US-09-221-017B-1095/c  
; Sequence 1095, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: ROSS, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PPI182  
; FILING DATE: 31-DEC-1997  
; PRIOR APPLICATION DATA: PPI1546  
; APPLICATION NUMBER: PPI1546  
; FILING DATE: 30-JAN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP2911  
; FILING DATE: 09-APR-1998  
; PRIOR APPLICATION DATA: PCT/AU98/01023  
; FILING DATE: 10-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monroy, Gladys H  
; REGISTRATION NUMBER: 32,430  
; REFERENCE/DOCKET NUMBER: 27340-20021.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-813-5600  
; TELEFAX: 650-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 1095:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3717 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: UNKNOWN  
; ORIGINAL SOURCE:  
; ORGANISM: PORPHYROMONAS GINGIVALIS  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1...3717

US-09-221-017B-1095

Query Match 6.0%; Score 79; DB 4; Length 3717;  
Best Local Similarity 56.3%; Pred. No. 7.8e-12;  
Matches 148; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
QY 605 GGGCTGTGTGGGAAGAGCAGGCGCCGTTTCTTCAGAGGGGTGGCGACTGCTGGTGAAGAG 664  
Db 2265 GTGATGGAAGGCAAGCATCGCCCGGACACTTCATATGGCTGGCGAGGTGGTAAGCAA 2206  
QY 665 TTGTTTAAATTTGTGGAGCCAGATGTGGCTGTGTTCGTTAAGAAGGATTATCAGCAGTGG 724  
Db 2205 CTCCTTCATGATGTGGAGCGCGACAAAGGCATATTTTCGGAGAAAAGGATTTCCAGCAATC 2146  
QY 725 CGGCTTATTCAGAGGATGTTTCGAGATCTTGATTTTTTCATAAAAGTGTAGTGTCTGAA 784  
Db 2145 GCCGTTATCCGAGCATGTTGAACCTGTGGCTTTACCGTTTACGATCGTGGCCTGTCT 2086  
QY 785 ATAACACGTTGATAATGATGGCTGGCAATGAGTTCACGTAATGTGCACCTTTTCACCTGAA 844  
Db 2085 ATCATCCGTTGAAGAGGACGGGTTTGCACTCAGTAGCCGCAACGTTCGCCCTCGGACCGAA 2026  
QY 845 GAGAGGAAAAGGCATCAAT 867  
Db 2025 GAGCGTGCCATAGCCCCCATCCAT 2003

RESULT 10  
US-09-318-794A-1  
; Sequence 1, Application US/09318794A  
; Patent No. 6177264  
; GENERAL INFORMATION:  
; APPLICANT: DEGUSSA AKTIENGESSELLSCHAFT  
; TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC  
; FILE REFERENCE: eggeling  
; CURRENT APPLICATION NUMBER: US/09/318,794A  
; PRIOR FILING DATE: 1999-05-26  
; PRIOR APPLICATION NUMBER: DE 198 55 312.9  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2164  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-318-794A-1

Query Match 5.4%; Score 70.8; DB 3; Length 2164;  
Best Local Similarity 55.8%; Pred. No. 1.1e-09;  
Matches 135; Conservative 0; Mismatches 107; Indels 0; Gaps 0;  
QY 620 AGCAGGCGCGGTTTCTTCAGAGGGGTGGCGACTGTGGTGACGAAAGTTGTTTAAATTTGTG 679  
Db 1523 AGCAGGCGCTGGCAATTTTCGATGGTGGCTACCGTGGTGGCGAAGCTGTTCAATTTGGTG 1582  
QY 680 GAGCCAGATGTGGCTGTTCGGTTAAGAAGATTATCAGCAGTGGCGGCTTATTCAGAGG 739  
Db 1583 CGCCCTGATCGTGCATATTTTGACAAAAGATGCTCAGCAGGTTCCGGTGATTCGGCGA 1642  
QY 740 ATGGTTCGAGATCTTGATTTTCCATAAAGTGTAGTGGTGAATAACACGTGATAAT 799  
Db 1643 TTGGTTGGCGATCTAGACATTCCTGGGAGATTGCTCCGATTTTCGCGGCC 1702  
QY 800 GATGGCTGGCAATGAGTTTCACTGTCACCTTTTCACCTGAAGAGAGGAAAAGGCA 859  
Db 1703 GATGGCTTAGCGGAATCCAGCCGCAATCAACGCTCTTTCTGCGGATCAGCGGCAAGCT 1762  
QY 860 CT 861  
Db 1763 CT 1764

RESULT 11  
US-09-318-793A-3

; Sequence 3, Application US/09318793A  
; Patent No. 6184007  
; GENERAL INFORMATION:  
; APPLICANT: Dusch, Nicole  
; APPLICANT: Kalinowski, Jörn  
; APPLICANT: Fuhrer, Alfred  
; TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC  
; FILE REFERENCE: 21123/260204  
; CURRENT APPLICATION NUMBER: US/09/318,793A  
; CURRENT FILING DATE: 1998-05-26  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2164  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (351)..(1163)  
; NAME/KEY: CDS  
; LOCATION: (1166)..(2002)  
US-09-318-793A-3

Query Match 5.4%; Score 70.8; DB 3; Length 2164;  
Best Local Similarity 55.8%; Pred. No. 1.1e-09;  
Matches 135; Conservative 0; Mismatches 107; Indels 0; Gaps 0;  
QY 620 AGCAGGCCGTTTCTTCAGAGGGTGGCGACTGTGGTACGAGTCTTTTAATATTGTG 679  
DB 1523 AGCAGGCCGTTTCTTCAGAGGGTGGCGACTGTGGTACGAGTCTTTTAATATTGTG 1582  
QY 680 GAGCCAGATGTGGTGTTCGTTAAGAGGATTATCAGCAGTGGCGGCTTATTCAGAGG 739  
DB 1583 GCCTGTATCGTGATATTTGACAAAAGATGCTCAGCAGTGGCGGTGATTCGGCGA 1642  
QY 740 ATGTTTCAGATCTTGATTTTCCATAAAGTAGTGTGTAATAACACGTGATAAT 799  
DB 1643 TTGGTTGGCATCTAGACATCCCGTGGAGATTCCGCCGTTCCGATTATTCGTGGCGCC 1702  
QY 800 GATGCCCTGGCAATGAGTTCAGTAAATGTGCACCTTTACCTGAAGAGAGGAAAGCA 859  
DB 1703 GATGCTTAGCCGAATCCAGCGCAATCAACGCTTTCTCGGATCAGCGAGCGCAAGT 1762  
QY 860 CT 861  
DB 1763 CT 1764

RESULT 12  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 5.3%; Score 69.2; DB 3; Length 4403765;  
Best Local Similarity 59.0%; Pred. No. 1.4e-07;  
Matches 138; Conservative 0; Mismatches 93; Indels 3; Gaps 1;  
QY 239 CGTCTGATGCGGGCCAGGCAAGCTCATTTGGGTGTGTCGCCACCATGGGCTTCTCCAC 298  
DB 4037206 CGGCACTGCGACTCACGGCCGCGAGTGTGTGGTCTACTATGGGTGCGCTGCAC 4037147  
QY 299 GCGGGCCACCTCTCGCTCTGTGGCCAGGCGCGCAACTCTCCG---ACGTGTCGCGCTC 355  
DB 4037146 GAAGGCCACCTCGGTTGGTGGTGGCGGCAAGCGGTGCGCGATCGTCTGCTGTG 4037087  
QY 356 TCCATCTACGTCAACCCGGGCAATTCCGCCCCACGAGGAGCTCTCCACTACCTCCCTCC 415  
DB 4037086 TCGATCTTCGTCAACCCGATGCAATTGGTGGCGGGAAGATCTCGAGGCTATCCCGC 4037027  
QY 416 GACTTCGACGGCGACGTAAAGAACTCGCTCGGTTCGCCGGCGCGCTCGACGTC 469  
DB 4037026 ACCCGGACGACGACCTGGGCAACTCGGGCCGAGCGGTGGAATCGCTTTC 4036973

RESULT 13  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 5.3%; Score 69.2; DB 3; Length 4411529;  
Best Local Similarity 59.0%; Pred. No. 1.4e-07;  
Matches 138; Conservative 0; Mismatches 93; Indels 3; Gaps 1;  
QY 239 CGTCTGATGCGGGCCAGGCAAGCTCATTTGGGTGTGTCGCCACCATGGGCTTCTCCAC 298  
DB 4045135 CGGCACTGCGACTCACGGCCGCGAGTGTGTGGTCTACTATGGGTGCGCTGCAC 4045076  
QY 299 GCGGGCCACCTCTCGCTCTGTGGCCAGGCGCGCAACTCTCCG---ACGTGTCGCGCTC 355  
DB 4045075 GAAGGCCACCTCGGTTGGTGGTGGCGGCAAGCGGTGCGCGATCGTCTGCTGTG 4045016  
QY 356 TCCATCTACGTCAACCCGGGCAATTCCGCCCCACGAGGAGCTCTCCACTACCTCCCTCC 415  
DB 4045015 TCGATCTTCGTCAACCCGATGCAATTGGTGGCGGGAAGATCTCGAGGCTATCCCGC 4044956  
QY 416 GACTTCGACGGCGACGTAAAGAACTCGCTCGGTTCGCCGGCGCGCTCGACGTC 469  
DB 4044955 ACCCGGACGACGACCTGGGCAACTCGGGCCGAGCGGTGGAATCGCTTTC 4044902  
RESULT 14  
US-09-249-585A-4  
; Sequence 4, Application US/09249585A  
; Patent No. 6417002  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:28:44 ; Search time 20 Seconds  
(without alignments)  
1490.615 Million cell updates/sec

Title: US-10-033-269-9  
 perfect score: 1588  
 Sequence: 1 MAPAPRVISDKASMRSHRS.....AWFGKVRLLDNMEINLSNV 310

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	927.5	58.4	313	2	T03924	probable pantoate-	
2	639.5	40.3	280	2	F97296	pantoate-beta-alan	
3	537	33.8	281	2	F97258	pantoate-beta-alan	
4	536	33.8	282	2	E84448	pantoate-beta-alan	
5	518.5	32.7	282	2	G70482	pantothenate synth	
6	516	32.5	283	2	E64736	pantoate-beta-alan	
7	514.5	32.4	285	2	B87517	pantoate-beta-alan	
8	514	32.4	283	2	A96446	pantothenate synth	
9	514	32.4	283	2	A85497	pantothenate synth	
10	513	32.3	283	2	H83860	pantothenate synth	
11	504.5	31.8	283	2	F90066	pantoate-beta-alan	
12	500	31.5	284	2	AG0524	pantoate-beta-alan	
13	499	31.4	293	2	H82303	pantoate-beta-alan	
14	498	31.4	284	2	AD0413	pantoate-beta-alan	
15	497	31.3	283	2	S55486	probable pantoate-	
16	497	31.3	513	2	S75604	pantothenate synth	
17	494	31.1	286	2	H69671	pantothenate synth	
18	489.5	30.8	278	2	G81148	pantoate-beta-alan	
19	488	30.7	281	2	F82832	pantoate-beta-alan	
20	487	30.7	283	2	G83055	pantoate-beta-alan	
21	480.5	30.3	278	2	F81874	probable pantoate-	
22	467	29.4	534	2	A12172	pantothenate synth	
23	465	29.3	292	2	A12984	pantoate-beta-alan	
24	465	29.3	345	2	F98298	pantoate-beta-alan	
25	457.5	28.8	285	2	A84953	pantoate-beta-alan	
26	456	28.7	281	2	G75430	pantoate-beta-alan	
27	455.5	28.7	279	2	T47120	pantoate-beta-alan	
28	451	28.4	345	2	S48389	probable membrane	
29	447	28.1	293	2	AC3451	pantoate-beta-alan	

## ALIGNMENTS

```

RESULT 1
T03924
probable pantoate-beta-alanine ligase (EC 6.3.2.1) - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C;Accession: T03924
R;Genschel, U.

```

```
Query Match      58.4%; Score 927.5; DB 2; Length 313;
Best Local Similarity 59.9%; Pred. No. 2.6e-65;
Matches 185; Conservative 51; Mismatches 60; Indels 13; Gaps 6;

QY 5 PRVISKAMRSRSMRAQCKLIGLVPTMGFLHAGLSIV-AQAQLSDVVA--VSII 60
       .:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
Db 7 PEVIRDKAAWRASRRRAEGKTAVVPTMGYLHQHLSLISAAAAAADPVAIVVTII 66
       .:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:
```

RESULT 2  
E72296  
pantoate-beta-alanine ligase - Thermotoga maritima (strain MSB8)



Db 275 N 275

RESULT 5  
G70482  
pantothenate synthetase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 29-Sep-1999  
C:Accession: G70482  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V.

Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: G70482  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-282 <AQF>  
A:Cross-references: GB:AR000774; NID:g2984324; PIDN:AA07847.1; PID:g2984326; GB:AE00065  
A:Experimental source: strain VP5  
C:Genetics:  
A:Gene: pnc  
C:Superfamily: pantoate-beta-alanine ligase

Query Match 32.7%; Score 518.5; DB 2; Length 282;  
Best Local Similarity 40.6%; Pred. No. 2.9e-33;  
Matches 123; Conservative 47; Mismatches 106; Indels 27; Gaps 5;

QY 5 PRVDSKASMSRSMRAQGLIGLVPTMGFLHAGHLSVAQAQRLSDVVAVSIYVNP 64  
Db 2 PLLFKIKLNLNFKLNKCEGKEIGFVPTMGYLGHRQLKLARMQNDIVVISIFNPT 61

QY 65 QFATEDLSTYPSDFDGVKKLASVPGGVVDVVFHPRNLYDYGNKGDDVAEAGMVSCVE 124  
Db 62 QFGEDEDYDRPDLERDLE--ICEEGVDVVFAPVDEIYPK----- 102

QY 125 SGSGHESWVRVEKLELGLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKDYQOWELIQ 184  
Db 103 ---GIRTKVCVGEGLKVLGEFFRGHFDGVATVVKLFNIVQPNRAYFGKDYQOLKIE 159

QY 185 RMVRDLDFSIVKIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLLRKSAAGDQV 244  
Db 160 QVVEDLNIPIVPIVREDGLAYSRNVHLSPEEREKALSINKSLLRKSAAGDQV 219

QY 245 HCEKLTNLV---IQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRLLDN 301  
Db 220 DAKRIEKAIRAFIERHPHVKG-VDYVEITDQ-NLNPKETVEKGRILLVAVRGNARLLDN 277

QY 302 MEI 304  
Db 278 WKV 280

RESULT 6  
E64736  
pantoate-beta-alanine ligase (EC 6.3.2.1) - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 03-Jun-2002  
C:Accession: E64736  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: E64736  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-283 <BLAT>  
A:Cross-references: GB:AE000122; GB:U00096; NID:g1786315; PIDN:AA073244.1; PID:g1786325;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: pnc  
C:Superfamily: pantoate-beta-alanine ligase

C:Keywords: coenzyme A biosynthesis; ligase

Query Match 32.5%; Score 516; DB 2; Length 283;  
Best Local Similarity 38.9%; Pred. No. 4.6e-33;  
Matches 118; Conservative 52; Mismatches 109; Indels 24; Gaps 3;

QY 7 VISKASMSRSMRAQGLIGLVPTMGFLHAGHLSVAQAQRLSDVVAVSIYVNP 66  
Db 3 IITLPLLRQIRRLRMEGRKVALVPTMGNLHGHMKLVDEAKARADVIVVVFVNP 62

QY 67 APTFEDLSTYPSDFDGVKKLASVPGGVVDVVFHP--RNLYDYGNKGDDVAEAGMVSCVE 124  
Db 63 DRPEDLARYPRTLQEDCEKLNK--RKVDLVFAFSPVKEIY-----P 100

QY 125 SGSGHESWVRVEKLELGLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKDYQOWELIQ 184  
Db 101 NGTETHYVDVPGVLTMLGASRPGHFRGVSTIVSKLFNLVQPDIAFCGKDFQQLALIR 160

QY 185 RMVRDLDFSIVKIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLLRKSAAGDQV 244  
Db 161 KVVADMGFDIEIVGVPIMRAKDGLALSSRNGYLTAEQRTKAPGLYKVLSSIAADKLQAGER 220

QY 245 HCEKLTNLV---IQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRLLDN 304  
Db 221 DLDEIITIAQOELNEKGFRADDIQRDADTLLEVSETSKRAVILVAWLGDARLLDN 280

QY 305 NLS 307  
Db 281 ELA 283

# RESULT 7

E87517  
pantoate-beta-alanine ligase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: E87517  
R:Nierman, W.C.; Feildlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: E87517  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <STO>  
A:Cross-references: GB:AB005673; NID:g13423661; PIDN:AAK24137.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2166  
C:Superfamily: pantoate-beta-alanine ligase

Query Match 32.4%; Score 514.5; DB 2; Length 285;  
Best Local Similarity 39.7%; Pred. No. 6.1e-33;  
Matches 122; Conservative 49; Mismatches 105; Indels 31; Gaps 6;

QY 4 APRVSDKASMSRSMRAQGLIGLVPTMGFLHAGHLSVAQAQRLSDVVAVSIYVNP 63  
Db 3 SPIIVTVAEHREHVRWAKAQRVAVPTMGALHGHLSVRLAQOQHAERVATVFNVP 62

QY 64 GQFAPTEDLSTYPSDFDGVKKLASVPGGVVDVVFHPRNLYDYGNKGDDVAEAGMVSCV 123  
Db 63 KQFAPHEPDFAYPRGEADAELALV--GCDLLFAPNATEMY----- 102

QY 124 ESGSGHESWVRVEKLELGLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKDYQOWELI 183  
Db 103 --APGFSTLVSVSGVSEPLEGAARQFFGVATVVAKLFIQSQADVAVFGKDYQOLQV 160

QY 184 QRMVRDLDFSIVKIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLLRKSAAGDQ 243  
Db 161 RRMARDLDIPVEIIGAPTARAEDGLALSSRNAYLSAEERAAVAALPTAKAAAAVAQGG 220

QY 244 VHCEKLTNLV---IQSVTDAG-GRIDYAEIVDQNNLEKVEQIKSPV-----VFCVAAWFGKVR 297

Db 221 -PIEDAERSAVALQAGFGQVDYVEIRASOLSRLG--PGPIGASGRILVAAMLGKTR 277  
 QY 298 LIDNMEI 304  
 Db 278 LIDNMAV 284

# RESULT 8

A99646

panthothenate synthetase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 08

C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C:Accession: A99646  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: A99646

A>Status: Preliminary  
 A:Molecule type: DNA

A:Residues: 1-283 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA033560.1; PID:gl3359593; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:

A:Gene: ECs0137  
 C:Superfamily: pantoate-beta-alanine ligase

Query Match 32.4%; Score 514; DB 2; Length 283;

Best Local Similarity 38.9%; Pred. No. 6.7e-33;

Matches 118; Conservative 52; Mismatches 109; Indels 24; Gaps 3;

QY 7 VISDKASMRSSRSNRAGCKLIGLVTMGFLHAGHLSLVAQARQLSDVAVSYVNPQGF 66

Db 3 IIETPLLRQIRRLRMGKRVALVPTMGNLHDGKMLVDEAKARADVIVNPNQF 62

QY 67 APTEDLSYPSDFDGVKVLASVPGVDVVFHP--RNLYDYKNGGGDVAAGGWSVCE 124

Db 63 DRPEDLARYPTLQEDCEKLNK--RKVDLVFAPSVEIY-----P 100

QY 125 SSGSHESWVRVEKLELGLCGKSRPVFFRGVATVTVTKLFNIVEPDVAVFGKDYQOWRLIQ 184

Db 101 NGTETHYVDVPGIETRMNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQV 160

QY 185 RMVRDLDFSIKVIKVIETRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQV 244

Db 161 KMWADMGFDIEIVGVPIMRAKDGALSSRNGVLTAEQKIAPIGLYKVLSSIAADKLQAGER 220

QY 245 HCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRILIDNMEI 304

Db 221 DLDEIIAAGQELNEKGFRADDIQIRDADTLLVSETSKRAVILVAALGDLARLIDNMV 280

QY 305 NLS 307

Db 281 ELA 283

# RESULT 9

A85497

panthothenate synthetase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: A85497  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A85497  
 A>Status: Preliminary

A:Molecule type: DNA  
 A:Residues: 1-283 <STO>

A:Cross-references: GB:AB005174; NID:gl2512848; PIDN:AAG54437.1; GSPDB:GN00145; UWGP:2  
 A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:  
 A:Gene: panC  
 C:Superfamily: pantoate-beta-alanine ligase

Query Match 32.4%; Score 514; DB 2; Length 283;

Best Local Similarity 38.9%; Pred. No. 6.7e-33;

Matches 118; Conservative 52; Mismatches 109; Indels 24; Gaps 3;

QY 7 VISDKASMRSSRSNRAGCKLIGLVTMGFLHAGHLSLVAQARQLSDVAVSYVNPQGF 66

Db 3 IIETPLLRQIRRLRMGKRVALVPTMGNLHDGKMLVDEAKARADVIVNPNQF 62

QY 67 APTEDLSYPSDFDGVKVLASVPGVDVVFHP--RNLYDYKNGGGDVAAGGWSVCE 124

Db 63 DRPEDLARYPTLQEDCEKLNK--RKVDLVFAPSVEIY-----P 100

QY 125 SSGSHESWVRVEKLELGLCGKSRPVFFRGVATVTVTKLFNIVEPDVAVFGKDYQOWRLIQ 184

Db 101 NGTETHYVDVPGIETRMNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQV 160

QY 185 RMVRDLDFSIKVIKVIETRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQV 244

Db 161 KMWADMGFDIEIVGVPIMRAKDGALSSRNGVLTAEQKIAPIGLYKVLSSIAADKLQAGER 220

QY 245 HCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRILIDNMEI 304

Db 221 DLDEIIAAGQELNEKGFRADDIQIRDADTLLVSETSKRAVILVAALGDLARLIDNMV 280

QY 305 NLS 307

Db 281 ELA 283

# RESULT 10

A83860

panthothenate synthetase panC [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: H83860  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi

Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans ar

A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: H83860

A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-283 <STO>  
 A:Cross-references: GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA05407.1; GSPDB:GN

A:Experimental source: strain C-125  
 C:Genetics:

A:Gene: panC  
 C:Superfamily: pantoate-beta-alanine ligase

Query Match 32.3%; Score 513; DB 2; Length 283;

Best Local Similarity 40.3%; Pred. No. 8e-33;

Matches 122; Conservative 46; Mismatches 95; Indels 40; Gaps 6;

QY 8 ISDKASMRSSRSNRAGCKLIGLVTMGFLHAGHLSLVAQARQLSDVAVSYVNPQGF 67

Db 7 ISD---LQALRIERBQKRSVGFVPTMGVLHGHLSLVKRAKEHDTVWMSIFVNPQFG 63

QY 68 PTEDLSYPSDFDGVKVLASVPGVDVVFHP--RNLYDYKNGGGDVAAGGWSVCE 127

Db 64 AGEDLDTYPRDFARD--ELAEAB--GVDILFY-----S 94

QY 128 GHESWVRVEKLEIAG-----LCGKSRPVFFRGVATVTVTKLFNIVEPDVAVFGKDYQOW 180

Db 95 TDENYFRPASVRLKVTQGVVDVLCASRPGHFGDGVTVVLKFLHVEFDAAVFGKDAQV 154

QY 181 RLQFORMVRDLDFSIKVIKVIETRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAG 240

Db 155 AVITMWEDNLNGVQIVPCATVREVDGLAKSSRNVRSEKBERKAPGLYQSLLAGREALD 214  
 Qy 241 DGQVHCEKLTNLVQSVTD-AGGRIDYAEIVDQNNLEKVEQIKGPVFCVAAMFGKVRLLI 299  
 Db 215 AGENDAIVIRIQSLERLTGRIDYEVLSYPRLOKIERIEETVILAVAYQFENARLLI 274  
 Qy 300 DNM 302  
 Db 275 DNL 277  
  
 RESULT 11  
 F90066  
 Pantoate beta-alanine ligase [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: F90066  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogun-  
 ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: F90066  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-283 <KUR>  
 A:Cross-references: GB:BAO00018; PID:g13702555; PIDN:BA843696.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: panc  
 C:Superfamily: pantoate-beta-alanine ligase  
  
 Query Match 31.8%; Score 504.5; DB 2; Length 283;  
 Best Local Similarity 37.2%; Pred No. 3.7e-32;  
 Matches 113; Conservative 58; Mismatches 102; Indels 31; Gaps 5;  
  
 Qy 6 RVISDKASMRSSMRSAQKGLIGLVTMGFLHAGHLSLVAQARQLSDVAVSYVNPQG 65  
 Db 3 KLITTVKEMQHI VKAAKRS GTTIGFITMGALHDGHLTWRESVSTNDITVSVFNPLQ 62  
 Qy 66 FAPTEDLSTYSPDPTGDVKKLASVPGGVVVVHPRNLVDYKKGGGDVABAGKVCVES 125  
 Db 63 FGPNEFDAYPRQIDKDLVLSEV--GADIVFHPAVEDIVFPELGIDIV-KVGPLADVLE- 118  
 Qy 126 GSGHESWVRVEKLEGLGICGKSRPFFFGVATVTKLENIVEPDVAFGKDYQWRLIQR 185  
 Db 119 -----GAKRPHFGGVTVVVKLFNIVMPDYAYFGKDAQQLAIVEQ 160  
 Qy 186 MVRDLDFSIVIGAEITRDNDGLAMSSRNVHLSPEEREKALSINKSLLRKSAAGDQGVH 245  
 Db 161 MVKDFNHAVEIIGDIVREADGLAKSSRNVLTEQEQEAHVLSKSLLLAQLYQDQ--- 217  
 Qy 246 CEKLTNLVQSVTD-----AGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAMFGKVRLLID 300  
 Db 218 -ERQSKIIDRVTEYLESHISGRIEEAVVSYQQLVEQHEITGRIFSLAVKFSKARLLID 276  
 Qy 301 NMEI 304  
 Db 277 NIII 280  
  
 RESULT 12  
 AG0524  
 Pantoate, beta-alanine ligase [imported] - Salmonella enterica subsp. enterica serovar Typh  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AG0524  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Comerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, R.  
A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* ser  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AGO524  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-284 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD01335.1; PID:g16501463; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY0199  
C:Superfamily: pantoate-beta-alanine ligase

Query Match 31.5%; Score 500; DB 2; Length 284;  
Best Local Similarity 37.4%; Pred. No. 8.4e-32;  
Matches 114; Conservative 57; Mismatches 106; Indels 28; Gaps 5;

Qy 7 VISDKASMRSSRSRWAOGKILGLVPTWTGFLHAGHLSLVAQAQLSDVVAVSIYNPGQP 66  
Db :  
3 IIEIPLQRHIRRUQEGKRVALVPTNGNLHDGHMKLVDEAKARADVIVISIFVNPMQF 62  
Qy 67 APTEDLSTPSDFDGDVKKLASVPGVDWTEHP--RNLYDYCKNGSGDVAEAGMVSCVE 124  
Db :  
63 DRPDDLVRYPITIQEDCEKLNK--RKVDYVFAPAEEIYPPQLEG----- 105  
Qy 125 SGSHESWRVREKLELGICGKSRRPVFRGVATVTTKLFNIPEPDVAVFGKDYQQWRLLIQ 184  
Db :  
106 -----QTVDVPGLSTMLEGRSRPHFRGSVTISKLENLIQPDIAACGEKDFQQLALIR 160  
Qy 185 RMVZDLDFSIXVGABITRDNDGLAMSSRNVLHSPEEREKALSINKSL--LRAKSAAGDG 242  
Db :  
161 KKVADMGGYDIENVGPPIRAKDXGHALSSRNAYLTAEQRKIAPGLYNVNNSIAEKLIAGNR 220  
Qy 243 QVHCKEKTLNVIOQSVTDAGGRIDYAEIIVDQNLEKVEQIKSPVFCVAAWFKGVRLIDNM 302  
Db :  
221 EL--QEIIATAEQELNEKGFRADDIQRINDADLTLETTSKRAVILAAAWLQARLIDNQ 278  
Qy 303 EINLS 307  
Db : : : : :  
279 SVTLA 283

RESULT 13  
H82303  
pantoate-beta-alanine ligase VC0591 [imported] - Vibrio cholerae (strain N16961 serogr  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: H82303  
R:Heidelberg, J., Emlava, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J  
chardson, D.; Jermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406. 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: H82303  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <HEI>  
A:Cross-references: GB:AE004144; GB:AE003952; NID:g9655017; PIDN:AAF93758.1; GSPDB:GN0  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor

A;map position: 1  
C;Superfamily: pantoate-beta-alanine ligase

	Query Match	31.4%;	Score 499;	DB 2;	Length 293;
	Best Local Similarity	37.0%;	Pred. No. 1e-31;		
	Matches 121;	Conservative 54;	Mismatches 80;	Indels 72;	Gaps 7;

QY   6 RVIDSKASMSRSRMRQAQKLCGLVPTMGFLHAGHLILVAQRQLSDTVAVSTIYNPGQ 65  
     | | | : | :  
Db   2 QVFADIAVLREQIKQIKREGRRVAFVPTMGNLHEGHTLVLRKARELDADVWVSIFINPMQ 61  
  
QY   66 FATEDSLTSPDFDGDVKKLASVPQGDVVFF--HPRNLVDYGKNGGGDVAAEGGMVSCV 123



```

Db      62 FRAEDKNYPTLEBLSKNG--EGVDLVTPPTETWYPOGLD----- 104
QY     124 ESGSGHESWVRVEKLELGLCGSRPVRFRGVATVTKLFNIVPEPDVAVFGKDYQOQWRLI 183
Db      105 -----KQTEVEVPEGLSYMLEGASRPGHFRGVATVTKLFNIVQPDVAVFGKDYQOQWRLI 159
QY     184 QRMVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSAAGDQV 243
Db      160 RQVDELDMDIEIVGATVRELDGLAMSSRNHLSPEEREKALSINKSLRAKSAAGDQV 215
QY     244 VHCEKLTNLVTSQVTDAGRIDYAEIVDQNNLEKVEQIKS-----PV- 285
Db      216 -----RGRDDYPSIIE---DAVDQLRAADLEPDEIFIRDARTLLPIS 255
QY     286 -----VFCVAWFGKVRLLIDNMEINL 306
Db      256 SESQQAIVLMSAFLGKVRLLIDNQVLDL 282

RESULT 14
AD0413
pantoate-beta-alanine ligase (EC 6.3.2.1) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 03-Jun-2002
C:Accession: AD0413
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitthead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0413
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <KUR>
A:Cross-references: GB:ALS90842; PIDN:CAC92632.1; PID:g15981329; GSPDB:GN00175
A:Gene: panc
C:Superfamily: pantoate-beta-alanine ligase
C:Keywords: coenzyme A biosynthesis; ligase

Query Match      31.4%; Score 498; DB 2; Length 284;
Best Local Similarity 39.7%; Pred. No. 1.2e-31;
Matches 120; Conservative 48; Mismatches 110; Indels 24; Gaps 6;

QY      7 VISDKASMSRSWRSRAQKLGIVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQV 66
Db      3 IIEITPLLRQRIWRQEQKRIALVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQV 62
QY     67 APTEDLSTYPSDFDGVKKLASVPGGVVVVHFRNLVDYKNGGGDVAAAGMVCVSESG 126
Db      63 ERPDLLAHYPRTLQEDCEKLTTR--HGADLVFAP-----AAADLYPAG-----LEK- 105
QY     127 SGHESWVRVEKLELGLCGSRPVRFRGVATVTKLFNIVPEPDVAVFGKDYQOQWRLI 186
Db     106 ---QTVYDVVPALSTTLLEGASRPGHFRGVATVTKLFNIVQPDVAVFGKDYQOQWRLI 162
QY     187 VRDLDFSIKVGAEITRDNDGLAMSSRNHLSPEEREKALSINKSL--LRKSAAGDQV 244
Db     163 VADMGYDINIVGPTVTRAKDGLALSSRNGVLTTEERQIAPQLSKINWALAEKMGALGERQI 222
QY     245 HCCKLTNLVTSQVTDAGRIDYAEIVDQNNLEKVEQIKSPVPVFCVAWFGKVRLLIDNMEI 304
Db     223 --DALLEEAAQLLRVGTPTDELFIIRDAETLQPTVDSQAVILMAALWGKARLLIDNQIV 280
QY     305 NL 306
Db     281 DL 282

```

RESULT 15  
S55486

```

probable pantoate-beta-alanine ligase (EC 6.3.2.1) - fission yeast (Schizosaccharomyce
N:Alternate names: protein ACSH10.08c
C:Species: Schizosaccharomyces pombe
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 03-Jun-2002
C:Accession: S55486; T38972
R:Connor, R.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55479
A:Accession: S55486
A:Molecule type: DNA
A:Residues: 1-283 <CON>
A:Cross-references: EMBL:Z49811; NID:g854599; PIDN:CAA89958.1; PID:g854607
R:Connor, R.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, May 1995
A:Reference number: Z21821
A:Accession: T38972
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-283 <CO2>
A:Cross-references: EMBL:Z49811; PIDN:CAA89958.1; GSPDB:GN00066; SPDB:SPAC5H10.08c
A:Experimental source: strain 972h-; cosmid CSH10
C:Genetics:
C:Gene: SPDB:SPAC5H10.08c
A:Map position: 1
C:Superfamily: pantoate-beta-alanine ligase
C:Keywords: coenzyme A biosynthesis; ligase

Query Match      31.3%; Score 497; DB 2; Length 283;
Best Local Similarity 40.1%; Pred. No. 1.4e-31;
Matches 122; Conservative 44; Mismatches 114; Indels 24; Gaps 3;

QY      6 RVISDKASMSRSWRSRAQKLGIVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQV 65
Db      2 QVLKEKLLIHQVQVDNWRKDCGNRIAFVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQV 61
QY     66 FAPTEDLSTYPSDFDGVKKLASVPGGVVVVHFRNLVDYKNGGGDVAAAGMVCV 123
Db      62 FNNPQDILLYPTMTDQDCSQQLNL--GVDLVYAPTVEELY----- 99
QY     124 ESGSGHESWVRVEKLELGLCGSRPVRFRGVATVTKLFNIVPEPDVAVFGKDYQOQWRLI 183
Db     100 PEGSQDITFVDVVKLSTMLEGASRPGHFRGVATVTKLFNIVPEPDVAVFGKDYQOQWRLI 159
QY     184 QRMVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSAAGDQV 243
Db     160 KRMVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSAAGDQV 219
QY     244 VHCEKLTNLVTSQVTDAGRIDYAEIVDQNNLEKVEQIKSPVPVFCVAWFGKVRLLIDNME 303
Db     220 GDLKLTAEITNLSTSRCFIPDQLEICDSTLTLEPFTAGTKNKNVILAAALWGKARLLIDNIQ 279
QY     304 INLS 307
Db     280 TTIN 283

```

Search completed: December 15, 2003, 13:34:17  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:29:19 ; Search time 17 Seconds  
(without alignments)  
857.546 Million cell updates/sec

Title: US-10-033-269-9  
Perfect score: 1588  
Sequence: 1 MAPAPRVISDKMSKRS .....AWFGKVLIDNMEINLSNV 310

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1286	81.0	307	1	PANC LOTJA
2	1101	69.3	310	1	PANC ARATH
3	927.5	58.4	313	1	PANC ORYZA
4	639.5	40.3	280	1	PANC THEMA
5	631.5	39.8	280	1	PANC THENE
6	545	34.3	283	1	PANC BRAJA
7	537	33.8	281	1	PANC CLOAB
8	536	33.8	282	1	PANC GAMGE
9	518.5	32.7	282	1	PANC AQUAE
10	516	32.5	283	1	PANC ECOLI
11	514.5	32.4	285	1	PANC CAUCR
12	514	32.4	283	1	PANC ECO57
13	513	32.3	283	1	PANC BACHD
14	506.5	31.9	281	1	PANC SHEON
15	504.5	31.8	283	1	PANC STRAM
16	504.5	31.8	283	1	PANC SPAAW
17	504	31.7	284	1	PANC SALT1
18	500	31.5	284	1	PANC SALT2
19	499	31.4	293	1	PANC VIBCH
20	498	31.4	284	1	PANC VERPE
21	497	31.3	283	1	PANC SCHPO
22	494	31.1	286	1	PANC BACSU
23	491	30.9	282	1	PANC SYN3
24	489.5	30.8	278	1	PANC NEIMB
25	488	30.7	281	1	PANC XYLEFA
26	487	30.7	283	1	PANC PSEAE
27	483	30.4	301	1	PANC VIBPA
28	482.5	30.4	286	1	PANC STAPB
29	480.5	30.3	278	1	PANC NEIMA
30	478	30.1	301	1	PANC VIBVU
31	475	29.9	291	1	PANC RHIME
32	472.5	29.8	280	1	PANC XANAC
33	467.5	29.4	283	1	PANC RHIL0

RESULT 1  
PANC LOTJA  
ID PANC LOTJA STANDARD; PRT; 307 AA.  
AC 024035;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pantoate--beta-alanine ligase precursor (EC 6.3.2.1) (Pantothenate synthetase) (Pantoate activating enzyme).  
GN PANC.  
OS Lotus japonicus.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.  
OX NCBI\_TaxID=34305;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 1-16, AND TISSUE SPECIFICITY.  
RC STRAIN=cv. Gifu / B-129; TISSUE=Root nodules;  
RX MEDLINE=99348031; PubMed=10417331;  
RA Genschel U., Powell C.A., Abell C., Smith A.G.;  
RT "The final step of pantothenate biosynthesis in higher plants: cloning and characterization of pantothenate synthetase from Lotus japonicus and Oryza sativum (rice).";  
RL Biochem. J. 341:669-678(1999).  
CC -1- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP + diphosphate + (R)-pantothenate.  
CC -1- PATHWAY: Pantothenate biosynthesis; last step.  
CC -1- SUBUNIT: Homodimer (Potential).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN LEAF AND ROOT.  
CC -1- MISCELLANEOUS: OPTIMAL PH IS 7.8. ACTIVITY DECREASES SHARPLY WITH INCREASING ACIDITY AND IS NIL AT PH 7. THERE IS ONLY A SLIGHT DECREASE TOWARDS HIGHER PH.  
CC -1- SIMILARITY: Belongs to the pantothenate synthetase family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Y10252; CAA71302.1; --  
InterPro; IPR003721; Pantoate\_ligase.  
DR Pfam; PF02569; Pantoate\_ligase; 1.  
DR TIGRFAMs; TIGR00018; pnc; 1.  
KW Pantothenate biosynthesis; Ligase.  
FT INIT.MET 0 0  
FT PROPEP 1 1  
FT CHAIN 2 307  
SQ SEQUENCE 307 AA; 34109 MW; 91C464EBE54E1642 CRC64;  
PANTOATE--BETA-ALANINE LIGASE.

Query Match 81.0%; Score 1286; DB 1; Length 307;  
Best Local Similarity 80.8%; Pred. No. 7.2e-99;  
Matches 248; Conservative 24; Mismatches 31; Indels 4; Gaps 2;

```

QY 4 APRV1SDKASMRSSMRRAQKGLIGLVTMGFLHAGHLSLVAQARQLSDVAVSIYVNP 63
Db 1 APW1SDKDEMRKWSRSMRSQKGLIALVPTMGFLHAGHLSLVRDAHNLADLVAVSIYVNP 60
QY 64 GQAPATEDISTYPSDFDGVKKLASVPGVDVVFHPRNLYDYGKNGGDDVAEAG--GMVS 121
Db 61 GQSPPTEDISAYPSDFDGLQKLMVPGVDVVFHPRNLYDYGGDGDVAEAGCGDGVVS 120
QY 122 CVE--SGSGHESWVRVEKLELGLCGKSRPVFFRGVATVTKLFNIVEPDVAVFGKKDYQ 179
Db 121 CVDRRSQFCHETWRAEKLKELCGKSRPVFFRGVATVTKLFNIVEPDVAVFGKKDYQ 180
QY 180 WRLIQRVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 239
Db 181 WKIIQRVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 240
QY 240 GDQVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAMFGKVR 299
Db 241 EDQIHCCKLNLVQSITEAGGRIDYAEIVDQNNLEKVEIKGPVFCVSAWFGKARLI 300
QY 300 DNMEINL 306
Db 301 DNIEINL 307

RESULT 2
PANC_ARATH STANDARD; PRT; 310 AA.
ID PANC_ARATH
AC Q9FKB3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN AT5G4840 OR K246.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RC MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RL DNA Res. 5:203-216(1998).";
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB012242; BAB09437.1;
CC InterPro: IPR003721; Pantoate_ligase.
CC Pfam: PF02569; Pantoate_ligase; 1.
CC TIGRFAMs: TIGR00018; panc; 1.
KW Pantothenate biosynthesis; Ligase.
SQ SEQUENCE 310 AA; 34137 MW; AD07D927625AD1A6 CRC64;

```

```

Query Match 69.3%; Score 1101; DB 1; Length 310;
Best Local Similarity 67.2%; Pred. No. 1.3e-83;
Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;
QY 5 PRIVISDKASMRSSMRRAQKGLIGLVTMGFLHAGHLSLVAQARQLSDVAVSIYVNP 64
Db 3 PEVIRDKDSMRKWSRSMRSQKGLIGLVTMGFLHAGHLSLVRQSLALTDVTVVSIYVNP 62
QY 65 QFAPATEDISTYPSDFDGVKKLASVPGVDVVFHPRNLYDYG-----KNGGDDVAEAG 118
Db 63 QFSPPTEDISTYPSDFDGLTKLAALSGGKVVVFNPKNLYDYGETKKINDGGN---GGR 119
QY 119 MVSCVSEGS--GHESWVRVEKLELGLCGKSRPVFFRGVATVTKLFNIVEPDVAVFGKKDY 177
Db 120 VVSCVEGGGLHETWIRVERLEKGFCKSRPVFFRGVATVTKLFNIVEPDVAVFGKKDY 179
QY 178 QWRLLQRMVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 237
Db 180 QWRLLQRMVRLDPSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 239
QY 238 AAGDQVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAMFGKVR 297
Db 240 SVAEKTKNCAELKDMIIQQVVGSRGVVTVIVDQTLGVEIKSGVVICVAAMFGKVR 299
QY 298 LIDNMEINLSM 308
Db 300 LIDNIEINLSL 310

RESULT 3
PANC_ORYZA STANDARD; PRT; 313 AA.
ID PANC_ORYZA
AC O24210;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN SEQUENCE FROM N.A.
RA MEDLINE=99348031; PubMed=10417331;
RA Genschel U., Powell C.A., Abell C., Smith A.G.;
RT "The final step of pantothenate biosynthesis in higher plants: cloning
RT and characterization of pantothenate synthetase from Lotus japonicus
RL Biochem. J. 341:669-678(1999).";
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y10253; CA471303.1;
CC PIR: T03924; T03924.
CC Grane; O24210;
CC InterPro: IPR003721; Pantoate_ligase.
CC Pfam: PF02569; Pantoate_ligase; 1.

```

```

DR   TIGRFAMs; TIGR00018; panC; 1.
KW   Pantothenate biosynthesis; Ligase.
SQ   SEQUENCE 313 AA; 33894 MW; 3074E74A49BF2C48 CRC64;

Query Match      58.4%; Score 927.5; DB 1; Length 313;
Best Local Similarity 59.9%; Pred. No. 2.8e-69;
Matches 185; Conservative 51; Mismatches 60; Indels 13; Gaps 6

Qy   5  PRVTSKASMSRWSRSRAOQKLGIVTMTGFLHAGHLSLV--AOARQLSDVVA--VSIT 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   7  PEVIRDAAMRAMSRRRRRAEKTVAVVTMTGYLHQHLSLISAAAAASADPVAIVVTI 66

Qy   61  VNPSQFAPTEDSLTPSDFDGVKVLASVPGVDVVFHPRNIY--DYKNGSGGDVAERAG 118
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   67  VNPSQFAPSEDLATPSDFAGDLRLAST-GWDAVENPPDLYVRGAGRRGAG----SGG 121

Qy   119  MVSCVE--SGSCHSESWRVEKLELGLCKQSRPVFPFGVATVTKLFNTEVPDPAVFGKD 176
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db   122  AISCLEAAGDGHETWVRERLEKGLCGASPVFPFGVATTVKULFNIEPDVPVFGKD 181

Qy   177  YQWRLLTQRMVRDLDFSIKVTGAEITRDNGLAMSSRNVLHSPEREKALSTINKSLRAK 236
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   182  YQWRVLTLPYWSGLDFGIEIMGSRNCARTDGLAMSRNVHLSRBEKKALSTIRSLVDAR 241

Qy   237  SAAGDGQVHCCKNLNLITQSVYTDAGGRIDYAEIVDONNLEKVEQIKSPVPCVAAWFGKV 296
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   242  TGALKGNVDSKQIKKKIKVQTUTETGGQVDYVEIVEQESLVPEQIDGPPWTCVAAWFGKV 301

Qy   297  RLIDNMEIN 305

Db   302  RLIDNIEID 310

```

```

RESULT 4
PANC_PANC
ID PANC THEME STANDARD; PRT; 280 AA.
AC Q9X036;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
DE PANC OR TM1077.
GN Thermotoga maritima.
OS Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OC NCBI_TaxID=2336;
OX [1]
RN RP SEQUENCE FROM N.A.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001768; AAD36154.1; -
CC DR

```

DR	PIR; E72296; E72296.
DR	TIGR; TM1077; --
DR	HAMAP; MF 00158; --
DR	InterPro; IPR004821; Cyt tran rel.
DR	InterPro; IPR003721; Pantoate_ligase.
DR	Pfam; PF02569; Pantoate_ligase; 1.
DR	TIGRFAMs; TIGR00125; cyt_tran_rel; 1.
DR	TIGRFAMs; TIGR00018; panC; 1.
DR	Pantothenate biosynthesis; Ligase; Complete proteome.
KW	SEQUENCE 280 AA; 32756 MW; 0BD1E4C572197163 CRC64;
QY	Query Match 40.3%; Score 639.5; DB 1; Length 280;
DB	Best Local Similarity 47.2%; Pred. No. 1.4e-45;
DB	Matches 142; Conservative 44; Mismatches 84; Indels 31; Gaps
QY	6 RVIDSKASMSRWSMRAGKLGIVPTMGFLHAGHLSVLAQARQLSDVVAISIVYVNPQQ 65
DB	2 RIITIEMKFFSEMRKKTIQFVPTMGVLHGHLSLVRARAENDVVVVSIFVNPQQ 61
QY	66 FAPTEDLSTYPSDFDGVGVKVLASVPGGVVDVVFHP--RMLY--DYGKNGSGDVAEAGMV 121
DB	62 FGPNEYERYPRDFEROKLEK--ENVDCIFHFSVEEMYPDF----- 103
QY	122 CVESGSHESWVRVEKLELGLGCGKSRPVFFRGVATVTKLFNIVPEPDVAVFGKKDYQQWR 181
DB	104 -----STVVEETKLSKHLCGRSRPGHFRGCVTVTKLFNIVKPHRAYFQKDAQQFR 155
QY	182 LIQWRDLDFSIKVI GAETRDNDGLAMSSRNVHLSPEEREKALSIKSLLRKASAGD 241
DB	156 VLRRVRLDNDVMEICPIVREPDGLAMSSRNVYLSPEERQALLSLYQSLKIAENLYLN 215
QY	242 GQVCEKLTNLVIOVSVDAGG-RIDYARIVQNNLEKVEQIKSPVFCVAAWFGKVRLLD 300
DB	216 GERDAEKIEMIKHLSRFDKVKIDYIVIEVDEETLEPVEKIDRKVIVAAWVGVARLLD 275
QY	301 N 301
DB	276 N 276
RESULT 5	
PANC_THENE	
ID	PANC_THENE STANDARD; PRT; 280 AA.
AC	O86953;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Pantoate--beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE	(Pantoate activating enzyme).
GN	PANC.
OS	Thermotoga neapolitana.
OC	Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX	NCBI_TaxID=2337;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=EZ2706-MC24;
RA	Zverlov V.V.;
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC	-I- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC	diphosphate + (R)-pantothenate.
CC	-I- PATHWAY: Pantothenate biosynthesis; last step.
CC	-I- SIMILARITY: Belongs to the pantothenate synthetase family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a> )
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
CC	-----
DR	EMBL; AJ007446; CAA07518.1; --
DR	HAMAP; MF 00158; --1.

```

DR InterPro: IPR004821; Cyt_tran_rel.
DR InterPro: IPR003721; Pantoate_ligase.
DR Pfam: PF02569; Pantoate_ligase; 1.
DR TIGRFAMs: TIGR000125; Cyt_tran_rel; 1.
DR TIGRFAMs: TIGR00018; panC; 1.
KW Pantothenate biosynthesis; Ligase.
SQ SEQUENCE 280 AA; 32608 MW; 169BEE73B5C5E075 CRC64;

Query Match 39.8%; Score 631.5; DB 1; Length 280;
Best Local Similarity 46.8%; Pred. No. 6.6e-45;
Matches 141; Conservative 44; Mismatches 85; Indels 31; Gaps 5;

QY 6 RVISDKASMSRSRAQKGLIGVPTMGFLHAGHLSLVAQAQRLSDVVAISIVYVNPQ 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 RIETTEMKKSEEMREKKTIGFVPTMGFLHAGHLSLVAQAQRLSDVVAISIVYVNPQ 61

QY 66 FAPTEDLSTYPSDFDGDVKKLASVPGGVVVVFP--RNLY--DYGKNGGGDVAAEGMWYS 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 FGNEDYERYPRDFERDKLEK--ENVDCVPHPSVEEMYPDP----- 103

QY 122 CVESGSHSWVRVEKLELGLCKSRPVFRGVATVVTKLFNIVEPDVAVFGKDYQQWR 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
104 -----STFVEETKSLKPLQGRSPGHFRGVCVTVTKLFNIVKPHRAYFGKDAQQFR 155

QY 182 LIQRMVRDLDFSIVKIGAEITRDNDGLAMSSRNHLSPEERKALSINKSLRAKSAAGD 241
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 VLRMVRDNDNDVEMTECPVREPDGLAMSSRNHLSPEERKALSINKSLRAKSAAGD 215

QY 242 GQVHCKELTNLVIQSVTDAGG-RIDYAEIVDQNNLEKVBQISPVVFCVAAVFGKRLID 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
216 GERDAVKIKEMINHLSDRDKIDYVEIVDETELEPVEKIDRKIVAAVWVGKRLID 275

QY 301 N 301
Db 276 N 276

RESULT 6
PANC_BRAJA STANDARD; PRT; 283 AA.
AC Q9AMR9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR BLR2102 OR ID912.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110SPC4;
RX MEDLINE=21101824; PubMed=11157954;
RA Goettfert M., Roethlisberger S., Kuendig C., Beck C., Marty R.,
RA Hennecke H.;
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb
RT DNA region of the Bradyrhizobium japonicum chromosome.";
RL J. Bacteriol. 183:1405-1412(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
RC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.

```

```

CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF322013; AAG61078.1; -.
DR EMBL: AF005942; BAC47367.1; -.
DR HAMAP: MF_00158; -. 1.
DR InterPro: IPR004821; Cyt_tran_rel.
DR InterPro: IPR003721; Pantoate_ligase.
DR Pfam: PF02569; Pantoate_ligase; 1.
DR TIGRFAMs: TIGR000125; Cyt_tran_rel; 1.
DR TIGRFAMs: TIGR00018; panC; 1.
KW Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 283 AA; 31322 MW; A2CA526734D9A607 CRC64;

Query Match 34.3%; Score 545; DB 1; Length 283;
Best Local Similarity 41.0%; Pred. No. 9.2e-38;
Matches 125; Conservative 48; Mismatches 104; Indels 28; Gaps 5;

QY 6 RVISDKASMSRSRAQKGLIGVPTMGFLHAGHLSLVAQAQRLSDVVAISIVYVNPQ 65
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 KVIYVAELRRALADVRNAEKRIQVPTMGFLHAGHLSLVAQAQRLSDVVAISIVYVNPQ 61

QY 66 FAPTEDLSTYPSDFDGDVKKLASVPGGVVVVFP--RNLYDYGKNGGGDVAAEGMWVSC 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 FGNEDLSRYPRDFARDEALCGS--AGVSIIFASAEIYP----- 100

QY 124 ESGSGHESWVRVEKLELGLCKSRPVFRGVATVVTKLFNIVEPDVAVFGKDYQQWR 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
101 ---AQSFSEFVPEGLAKPLCGAFRPHFRGVATVVTCLFNVMQPDVAVFGKDYQQCAVI 157

QY 184 QRMVRDLDFSIVKIGAEITRDNDGLAMSSRNHLSPEERKALSINKSLRAKSAAGDQ 243
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 RMTVDNLNLPLEIVTVTVTRFEDGLAMSSRNHLSPEERKALSINKSLRAKSAAGDQ 217

QY 244 VHCCKELTNLVIQSVTDAGG-RIDYAEIVDQNNLEKVBQISPVVFCVAAVFGKRLIDNM 302
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 RDAATLALARRHL-ERVDRLQYLELVDPGTLRIADSPFLRYPAVLCAVAVGSTRLLDNV 276

QY 303 EINLS 307
Db : : : : :
277 VLWS 281

RESULT 7
PANC_CLOAB STANDARD; PRT; 281 AA.
AC Q97F38;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR CAC2915.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;

```





```

PANC CAUCR
ID PANC CAUCR STANDARD; PRT; 285 AA.
AC Q9A6C8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantoate--beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR CC2166.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA DeBoy R.T., Dodson R.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Haft D.H.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005888; AAK24137.1; -.
DR PIR; E87517; E87517.
DR HAMAP; MF 00158; -.
DR InterPro; IPR004821; Cyt tran_rel.
DR InterPro; IPR003721; Pantoate_ligase.
DR Pfam; PF02569; Pantoate_ligase; 1.
DR TIGRFAMs; TIGR00125; cyt_tran_rel; 1.
DR TIGRFAMs; TIGR00018; panC; 1.
DR Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 285 AA; 30052 MW; C5B89BCD51D8DFA8 CRC64;

Query Match 32.4%; Score 514.5; DB 1; Length 285;
Best Local Similarity 39.7%; Pred. No. 3.1e-35;
Matches 122; Conservative 49; Mismatches 105; Indels 31; Gaps 6;

QY 4 APRVISDKASRSRSRAGKLGILVPTGFLHAGLSVAQARQLSDVVAIVNVP 63
DB 3 SPIIVRTVAENRHRVRAAGQAVVPTTGALHAGLSVAQARQLSDVVAIVNVP 62
QY 64 GQFAPTEDLSTYDFDGDVKKLASVPGGVDDVWFPHRNLYDYKNGGDDVAEAGGWYSCV 123
DB 63 KQFAPHEFDYPRGEADAERKALV--GCDLLFAPNATENY----- 102
QY 124 ESGSGHSRWKVEKLEGLCKSRPFRPGVATVTKLFNIVEPDVAVFGKKDYQOWRL 183
DB 103 --AFGFTSLVSGVSEPLEGARFPQFGGVATVVKLFQSQADVAVFGKDYQQLQV 160
QY 184 QRMVRLDLSKIVTGAETIRNDGLAMSRNVHLSPEEREKALSKLSLRKSAAGDQ 243
DB 161 RRMARDLDPVEIIGAPTAREADGLALSSRNAYLSABERAAVALPTAMRAAAAAVQGG 220
QY 244 VHCKLTLNLVQSTVDAG-GRIDVAETVDQNNLEKVEIKSPV-----VFCVAWFQKVR 297

```

```

221 -PIEDAERSVAALQAGFGQVDYVEIREASDLSRLG--PGPIGEASGRILVAALWGKTR 277
298 LIDNMEI 304
278 LIDNNAV 284

RESULT 12
PANC_ECO57
ID PANC_ECO57 STANDARD; PRT; 283 AA.
AC Q8X930;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantoate--beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR Z0144 OR ECS0137.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005189; AAG54437.1; -.
DR PIR; AP002550; BAB33560.1; -.
DR PIR; A95497; A85497.
DR PIR; A95497; A95497.
DR HAMAP; MF 00158; -.
DR InterPro; IPR004821; Cyt tran_rel.
DR InterPro; IPR003721; Pantoate_ligase.
DR Pfam; PF02569; Pantoate_ligase; 1.
DR TIGRFAMs; TIGR00125; cyt_tran_rel; 1.
DR TIGRFAMs; TIGR00018; panC; 1.
DR Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 283 AA; 31567 MW; 575364697ADAFB20 CRC64;

Query Match 32.4%; Score 514; DB 1; Length 283;
Best Local Similarity 38.9%; Pred. No. 3.3e-35;
Matches 118; Conservative 52; Mismatches 109; Indels 24; Gaps 3;

```



```

QY 7 VISDKASMRWSRSMRAQKGLIGLVPTMGFLHAGHLSLVAQARQLSDVVAVSIVNPGQF 66
D 3 IIBTLLPQRIIRLRMEGRKVALVPTMGHGLVDEAKARADVIVVIFVNPMPQF 52
QY 67 APTEDLSTPSPDGDVKKLASVPGVDVVVPH--RNLYDYKNGKGGDVAEAGGMVSCVE 124
D 63 DRPEDLARYPTLQEDCEKLNK--RKVDLVFAPSVKEIY-----P 100
QY 125 SGSGHESVYRVEKLELGLCKSRPVFFRGVATVTKLFNIVPDVAVFGKDYQOORLQ 184
D 101 NGTEHTYVDVPGSLTMEGASRPHGRGVSTIVSKLFLNLVQDIACFGEKQFQALIR 160
QY 185 RMYRDLDFSIKIVGABITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDGV 244
D 161 KXVADMGFDIEIVGPFIMRAKQGLALSSRNGYLTAQRIAPGLYKVLSSIAKQAGER 220
QY 245 HCEKLTNLVQSVDAGGIDYAEIVDQNNLEKVEQIKSPVFCVAAMFGKVRLLDNMI 304
D 221 DDEIIIAIAGQELNEKGFRADDIQIRDADTLLEVSSETSRAVILVAALGLDARLIDNKW 280
QY 305 NLS 307
D 281 ELA 283

RESULT 13
PANC BACHD
ID PANC BACHD STANDARD; PRT; 283 AA.
AC O9KCB;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantoate--beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (pantoate activating enzyme).
GN PANC OR BH1688.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001512; BAB05407.1; -.
CC PIR; H83860; H83860.
CC
CC HAMAP; MF 00158; -.
CC
CC InterPro; IPR004921; Cyt tran_rel.
CC
CC InterPro; IPR003721; Pantoate_ligase.
CC
CC Pfam; PF02569; Pantoate_ligase; 1.
CC
CC TIGRFAMs; TIGR00125; cyt tran_rel; 1.
CC
CC TIGRFAMs; TIGR00018; panC; 1.
CC
CC Pantothenate biosynthesis; Ligase; Complete proteome.
CC
CC SEQUENCE 283 AA; 31703 MW; 005898061568E4BC3 CRC64;
DR
DR HAMAP; MF 00158; -.
DR
DR InterPro; IPR004921; Cyt tran_rel.
DR
DR InterPro; IPR003721; Pantoate_ligase.
DR
DR Pfam; PF02569; Pantoate_ligase; 1.
DR
DR TIGRFAMs; TIGR00125; cyt tran_rel; 1.
DR
DR TIGRFAMs; TIGR00018; panC; 1.
DR
DR Pantothenate biosynthesis; Ligase; Complete proteome.
DR
DR SEQUENCE 283 AA; 31703 MW; 005898061568E4BC3 CRC64;
DR
DR TIGR; SO0869; -.

Query Match 32.3%; Score 513; DB 1; Length 283;
Best Local Similarity 40.3%; Pred No. 4e-35; Mismatches 46; Indels 40; Gaps 6;
Matches 122; Conservative 46; Mismatches 95; Indels 40; Gaps 6;

QY 8 ISDKASMRWSRSMRAQKGLIGLVPTMGFLHAGHLSLVAQARQLSDVVAVSIVNPGQF 67
D 7 ISD---LQALRIEREQKRSVGVPTMGYLGHLHLVKAKEEHDVTVMISFVNPLOFG 63
QY 68 PFEDLSTPSPDGDVKKLASVPGVDVVVPH--RNLYDYKNGKGGDVAEAGGMVSCVSGS 127
D 64 AGEDLDTTTPDFARD-EQLAEAB-GVDILFYP-----S 94
QY 128 GHESVYRVEKLELG-----LCGSRPVFFRGVATVTKLFNIVPDVAVFGKDYQOOW 180
D 95 TDEWYPRASVRLKVTQGVVLCGASRPHGRGVSTIVSKLFLNLVQDIACFGEKQFQALIR 154
QY 181 RMYRDLDFSIKIVGABITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAG 240
D 155 AVITNMVEDLNVGVQVPCATVREVVDGLAKSRNVLSSEKKEAPGLYQSLLAGREALD 214
QY 241 DGVHCEKLTNLVQSVD--AGRIDYAEIVDQNNLEKVEQIKSPVFCVAAMFGKVRLLI 299
D 215 AGEKDAVIRERIRQSLERLGRIDYVEVLSYPRLOKIERIETVILAVAYQFENARLI 274
QY 300 DNM 302
D 275 DNL 277

RESULT 14
PANC SHEON
ID PANC SHEON STANDARD; PRT; 281 AA.
AC O8EJHO;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pantoate--beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR SO0869.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Dougherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vanatavehan J., Weidman J., Imprial M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feildblum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis";
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE015532; AAN53945.1; -.
CC TIGR; SO0869; -.

```

Search completed: December 15, 2003, 13:34:47  
Job time : 18 secs



Db 3 PEVIRKDSMRKWSRAMRSQKTI GLVPTMGVYLHEGHSLSVRQSLALTDVTVVSVYVNPQ 62  
QY 65 QPAPTEDLSTYPSDFDGVKKLASVPGVDVVFPHPRNLVDYD-----KNGGGDVAEAGG 118  
Db 63 QFSPTEDLSTYPSDFSGDLTKLAALSGGKVVVFVFNPKNLYDYGETKKINDGGN---GGR 119  
QY 119 MVSCEVSGS-GHESVVRVEKLELGLCGKSRPVFFRGVATVTKLNIPEVDVAVFGKDY 177  
Db 120 VVSCVEEGGIGHETWIRVERLEKGLCGKSRPVFFRGVATVTKLNIPEVDVAVFGKDY 179  
QY 178 QQWRLIQRVRLDLSFKIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKS 237  
Db 180 QQWRLIQRVRLDLSFKIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKS 239  
QY 238 AAGDGVHCEKTLNVLQSVTDAGSDIYAEIVDQNNLEKVPQIKSPVVFCAAWFGKVR 297  
Db 240 SVAEGKTNCAELKMI IQQVGSAGRDVYVEIVDQETLGVBEIKSGVVICVAWFGTVR 299  
QY 298 LIDNMEINLSM 308  
Db 300 LIDNMEINLSL 310

## RESULT 2

Q8GDP0 PRELIMINARY; PRT; 315 AA.  
AC Q8GDP0;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Fragment).  
OS Helicobacillus mobilis.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;  
OC Helicobacillus.  
OX NCBI\_TaxID=28064;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22337798; PubMed=12446909;  
RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,  
RA Blankenship R.E.;  
RT "Whole-genome analysis of photosynthetic prokaryotes";  
RL Science 298:1616-1620(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Liolios K.G., Chu L., Ostrovskaya O., Mendiyaeva N., Koukharenko V.,  
RA Gerdes S., Kyripides N., Overbeek R.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY142936; AAN87540.1; --  
KW Ligase.  
FT NON\_TER  
ET 315  
SQ SEQUENCE 315 AA; 35339 MW; DBA58CC255F8F40C CRC64;

Query Match 38.8%; Score 616.5; DB 2; Length 315;  
Best Local Similarity 42.6%; Pred. No. 9.6e-44;  
Matches 132; Conservative 57; Mismatches 94; Indels 27; Gaps 5;

QY 3 PAPRIVDSKMSRWSRMRAGKLI GLVPTMGFLHAGHSLSVAQRLSDVVAIVSYVNPQ 62  
Db 25 PYMRLISSVNMMAWAKEQRRVGHGTIGLVPTMGYLHEGHLTLMRRAKENCERKVVVSIFN 84  
QY 63 PQGFAPTEDLSTYPSDFDGVKKLASVPGVDVVFPHPRNLVDYD-----KNGGGDVAEAGG 120  
Db 85 PLOFGNGEYEPRLDTRD-SQAD-SAGVDVIFAPVKMYP----- 126  
QY 121 SCVESGSHESVVRVEKLELGLCGKSRPVFFRGVATVTKLNIPEVDVAVFGKDYQOW 180  
Db 127 -----KGYSSFVEQVSDHLGCAAPGHFGVTVVSKLNIPEVDVAVFGKDYQOW 180  
QY 181 RLIIQRVRLDLSFKIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKS 240  
Db 181 AIIRRMVEDLNMGIAIVGPIVREADGLALSSRNHLSPEEREKALSINKSLRAKS 240  
QY 241 DQGVHCEKTLNVLQSV-TDAGSDIYAEIVDQNNLEKVPQIKSPVVFCAAWFGKVR 299

Db 241 EGEAEAIROBIKIVIEAEPLANIDYVQIVDNRFIQPVREGECLIALAVRFGKTRLI 300  
QY 300 DNMEINLSMN 309  
Db 301 DNLVMEVSPN 310  
RESULT 3  
Q9AMR9 PRELIMINARY; PRT; 283 AA.  
AC Q9AMR9;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE ID912.  
GN ID912.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=110spc4;  
RX MEDLINE=21101824; PubMed=11157954;  
RA Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R.,  
RA Hennecke H.;  
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb  
RL J. Bacteriol. 183:1405-1412(2001).  
DR EMBL; AF322013; AAG61078.1; --  
DR InterPro; IPR004821; Cyt tran rel.  
DR InterPro; IPR003721; Pantoate ligase.  
DR Pfam; PF02569; Pantoate ligase; 1.  
DR TIGRFAMs; TIGR00125; cyt tran\_rel; 1.  
DR TIGRFAMs; TIGR00018; panC; 1.  
SQ SEQUENCE 283 AA; 31322 MW; A2CA526734D9A607 CRC64;  
Query Match 34.3%; Score 545; DB 2; Length 283;  
Best Local Similarity 41.0%; Pred. No. 8.8e-38;  
Matches 125; Conservative 48; Mismatches 104; Indels 28; Gaps 5;  
QY 6 RVISDKASMRWSRMRAGKLI GLVPTMGFLHAGHSLSVAQRLSDVVAIVSYVNPQ 65  
Db 2 KVITKVAELRRALADVRAEAKRI GIVPTMGVYLHAGHSLSVAQRLSDVVAIVSYVNPQ 61  
QY 66 FAPTEDLSTYPSDFDGVKKLASVPGVDVVFPHPRNLVDYD-----KNGGGDVAEAGG 123  
Db 62 FGPNEILSRYPDFARDEALCGS--AGVSIIFAPSAEIIYP----- 100  
QY 124 ESGSGHESVVRVEKLELGLCGKSRPVFFRGVATVTKLNIPEVDVAVFGKDYQOWRLI 183  
Db 101 --AQFESFVPPGELAKPLCGAFRGHFRGVATVTKLNNVQPDVAVFGKDYQOCVAVI 157  
QY 184 QMVRDLDFSIKIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKS 243  
Db 158 RMTVDLMLPIETVETVTRPFDGLAMSSRNHLSPEEREKALSINKSLRAKS 217  
QY 244 VHEKLTNLVLIQSVTDAGSDIYAEIVDQNNLEKVEQ-IKSPVVFCAAWFGKVR 302  
Db 218 RDAATLIALARRHL-ERVDRLQYLELVDPTGLRIADSPRLYPVLCVAAVVGSTRLDNV 276  
QY 303 EINLS 307  
Db 277 VLSWS 281  
RESULT 4  
Q8EIH0 PRELIMINARY; PRT; 281 AA.  
AC Q8EIH0;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pantoate-beta-alanine ligase.
GN PANC OR SO0869.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Meyer T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Read T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Umayam L.A., White O., Wolf A.M.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AB015532; AAN53945.1; -.
DR TIGR; SO0869; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 281 AA; 30435 MW; CC7526CCB236BD3C CRC64;

Query Match 31.9%; Score 506.5; DB 16; Length 281;
Best Local Similarity 39.0%; Pred. No. 1.5e-34;
Matches 114; Conservative 48; Mismatches 109; Indels 21; Gaps 3;

QY 14 MRSWSRMAQCKLGLVPTMGFLHAGLSVAQARQLSDVAVSYNPGQFAPTELDLS 73
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
10 IRTQVRARAKGETVAFVPTMGNLHQHITLKEAAKCDHVASIFVNPQFQGNEDLD 69
QY 74 TYPSPDFDGDVKKLASVPGGVVDFHPRNLYDYGKGGDVAEAGGMVSGSGHESWV 133
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
70 AYPTLEADSQLTA--AGALLFTPTPAITYPK-----GLAQQTIV 109
QY 134 RVEKLELGCGSRPFRFGVATVTKLFNIVEPDVAVFGKKDYQOQRLIQRMVRLDIFS 193
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
110 EVPGISDVLGCASRPGHFGVATVCKLFNIVQPDIAFFGNKYQQLLVIRTWEDLSLP 169
QY 194 IKVIGAEITRNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQGVHCKRLNLV 253
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
170 IEIIGDITREASGLAMSRNGYLTAQEAAPALKKAIDAMAQIKQ3-ISIEQVTEEA 228
QY 254 IQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVVFCAAMFGKVRLLDNNWEIN 305
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
229 KASITAGFTPDYLEVRHADTLAKAETQDKALVILAAAYLGAARLIDNLRFD 280

RESULT 5
Q8DG73 PRELIMINARY; PRT; 513 AA.
AC Q8DG73;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Pantoate synthetase / cytidylate kinase.
GN PANC OR TLL2450.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium

```

```

RT Thermosynechococcus elongatus BP-1.;
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005377; BAC10002.1; -.
KW Kinase; Complete proteome.
SQ SEQUENCE 513 AA; 56833 MW; 259DDDF0476613222 CRC64;

Query Match 31.9%; Score 506; DB 16; Length 513;
Best Local Similarity 41.6%; Pred. No. 4.1e-34;
Matches 119; Conservative 42; Mismatches 83; Indels 42; Gaps 6;

QY 28 IGLVPTMGFLHAGLSVAQARQLSDVAVSYNPGQFAPTELDLSYPSDFDGVKKLA 87
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
26 VGFVPTMGALGGHAALIRARQCDVVVVSIFVNPQFQGBDLERYPRALADTALCQ 85
QY 88 SVPGGVVDFHPRNLYDYGKGGDVAEAGGMVSGSGHESWVR-VEKLELGCGKS 146
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
86 QL--GVDLLFVP-----SVAE-----LYPTGMKGLTVPEPRTELHLCGRS 125
QY 147 RPVFRGVATVTKLFNIVEPDVAVFGKKDYQOQRLIQRMVRLDIFSIKVIGAITRND 206
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
126 RPHFRGVATVTKLLHLVQPDRAVFGQKDAQQLAIRRCVADLNLVDVEIIGCPVRDAD 185
QY 207 GLAMSRNVHLSPEEREKALSINKSLRAKSAAGD-----QVHCEKLTNLVIQ 255
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
186 GLASSRNQYLSAEERATALLSLSLEVAQAARFGQGWASLLLAQVQDHLRQFPQL--- 242
QY 256 SVTDAGGRIDYAEIVDQNNLEKVEQIKSPVVFCAAMFGKVRLLDN 301
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
243 -----RLDYAEIVHPTQLVPLRIETVGLLAIAGWVGQTRLIDN 281

RESULT 6
Q8F394 PRELIMINARY; PRT; 285 AA.
AC Q8F394;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Panthothenate synthetase [EC 6.3.2.1].
GN PANC OR LA2514.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011418; AAN49713.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 285 AA; 32087 MW; 486E5934B40BA4A2 CRC64;

Query Match 30.7%; Score 488; DB 16; Length 285;
Best Local Similarity 37.8%; Pred. No. 5.7e-33;
Matches 115; Conservative 54; Mismatches 105; Indels 30; Gaps 7;

QY 5 PRIVISDKASMRWSRSRAQKGLIGLVPTMGFLHAGLSVAQARQLSDVAVSYNPG 64
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7 PEEVLDQ--VRLW---KAQKRLGFTVPTMGILHGHASLFEECISKADKTVVSIFVNP 60
QY 65 QFAPTELDLSYPSDFDGDVKKLASVPGGVVDFHPRNLYDYGKGGDVAEAGGMVSCVE 124
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 QFNDPEDYAKYPVNTGDKLCSKK--VDLVFLPNKETIY-PDGIPDIV----- 107
QY 125 SSGSHESWVRVEKLELGCGSRPFRFGVATVTKLFNIVEPDVAVFGKKDYQOQRLIQ 184
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
108 -----LKIPNLKMSLCAVSRPGHFGVLLVIFRLHFVQDPFAFFGKKDYQYLLIR 159
QY 185 RMVRDLDFSITKVGAEITRNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGDQV 244
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
160 EFCNTLAFPIEVIGCTVRSQQGLASSRNLSETEKEESLTIYRSLKLGNGQIFSG-I 218
QY 245 HCEKLTNLVIQSVTDAGG--RIDYAEIVDQNNLEKVEQIKSPVVFCAAMFGKVRLLDN 302

```











GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:32:29 ; Search time 41 Seconds  
(without alignments)  
1200.128 Million cell updates/sec

Title: US-10-033-269-9  
Perfect score: 1588  
Sequence: 1 MAPAPRVISDKASRWSRS.....AWFGKVLIDNNEINLSMNV 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.\*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1286	81.0	308	20 AAY42402	Amino acid sequenc
2	1101	69.3	310	23 ABB93747	Herbicidally activ
3	1101	69.3	643	21 AAG39461	Arabidopsis thalia
4	1101	69.3	653	21 AAG39460	Arabidopsis thalia
5	1101	69.3	654	21 AAG39459	Arabidopsis thalia
6	1098	69.1	310	21 AAG28281	Arabidopsis thalia
7	1098	69.1	322	21 AAG28280	Arabidopsis thalia
8	1071	67.4	299	21 AAG28282	Arabidopsis thalia
9	927.5	58.4	333	20 AAY42403	Amino acid sequenc

10	494	31.1	286	22	AAU01245	B. subtilis pantot
11	482.5	30.4	288	23	ABE39476	Staphylococcus epi
12	455.5	28.7	279	21	AAE10683	C. glutamicum panc
13	455.5	28.7	279	21	AAE10032	C. glutamicum panc
14	455.5	28.7	279	21	AAE10034	C. glutamicum panc
15	455.5	28.7	279	22	AAG89878	C. glutamicum prote
16	455.5	28.7	279	22	AA879940	Corynebacterium gl
17	454.5	28.6	259	24	ABF78440	N. gonorrhoeae ami
18	427	26.9	278	24	AAG79677	Human ENZM-7, incy
19	418	26.3	285	23	ABB48007	Listeria monocytog
20	405	25.5	275	23	ABU51388	Helicobacter pylor
21	372.5	23.5	229	22	AG822380	S. epidermidis ope
22	357	22.5	481	22	ABG28634	Novel human diagno
23	332	20.9	903	22	ABG22536	Novel human diagno
24	332	20.9	939	22	ABG25533	Novel human diagno
25	303.5	19.1	98	21	AAG55083	Arabidopsis thalia
26	230	14.5	111	23	ABU51871	Helicobacter pylor
27	228	14.4	266	22	AG822701	C. glutamicum prote
28	228	14.4	266	22	AA892701	Corynebacterium gl
29	160.5	10.1	117	23	ABU50860	Helicobacter pylor
30	103.5	6.5	421	18	AAW19735	Sugar biosynthesis
31	103.5	6.5	421	18	AAW19735	S. erythraea desosa
32	97.5	6.1	1297	23	AAU96927	Haemophilus influe
33	97.5	6.1	1298	23	AAU96928	Haemophilus influe
34	94.5	6.0	1297	23	AAU96930	Haemophilus influe
35	94	5.9	659	24	ABU26131	Aspergillus fumiga
36	94	5.9	975	20	AA42711	Rice GI protein (p
37	93.5	5.9	423	24	ABJ18791	Pseudomonas aerugi
38	93	5.9	422	22	AB882216	Megalomicin biosyn
39	92	5.8	332	23	ABP27548	Streptococcus poly
40	91.5	5.8	657	22	AAU31467	Novel human secret
41	91	5.7	339	22	AAG90770	C. glutamicum prote
42	91	5.7	339	22	AA876545	Corynebacterium gl
43	90.5	5.7	657	13	AA829580	FMR-1 gene product
44	90.5	5.7	997	22	ABG29047	Novel human diagno
45	90.5	5.7	1118	22	ABG24324	Novel human diagno

#### ALIGNMENTS

RESULT 1  
AAY42402  
ID AAY42402 standard; Protein; 308 AA.  
XX AC AAY42402;  
XX AC  
XX 02-DEC-1999 (first entry)  
XX  
XX Amino acid sequence of L. japonicus Pantothenate Synthetase (PS).  
XX pantothenate synthetase; cloning; herbicide; biosynthesis;  
XX Coenzyme A precursor; pantoate-beta-alanine ligase.  
XX  
XX Lotus japonicus.  
XX  
XX WO9942565-A1.  
XX  
XX 26-AUG-1999.  
XX  
XX 02-JUN-1998; 98WO-EP03261.  
XX  
XX 31-MAY-1997; 97GB-0011163.  
XX 27-JUN-1997; 97GB-0013477.  
XX  
XX (AGRE ) HOECHST-SCHERING AGREVO GMBH.  
XX Abell C, Smith AG, Genschel U, Laber B;  
XX WPI; 1999-527466/44.  
XX N-PSDB; AAZ20926.  
XX  
XX New isolated DNA molecule, useful in herbicidal compositions -

XX PS Claim 3; Fig 1.2; 64pp; English.

XX CC This is the amino acid sequence of the Lotus japonicus Pantothenate

XX CC synthetase (PS) enzyme, which is essential in the biosynthesis of

XX CC Coenzyme A.

XX CC The invention provides methods to assay proteins for PS activity, and

XX CC the ability of compounds to inhibit PS.

XX CC The assay compounds which are found to inhibit PS can be used as

XX CC herbicides.

XX CC The nucleotide probe can be used to amplify PS coding sequences from a

XX CC chosen organism via the process of polymerase chain reaction.

XX CC A number of assays have previously been used to measure PS activity

XX CC but are unsuitable for large screen biochemical screening of compounds

XX CC to find PS inhibitors. In this invention however the isolation of the

XX CC DNA molecule encoding PS enables a host cell to produce the protein

XX CC which can then be tested against compounds potentially able to inhibit

XX CC the enzyme.

XX SQ Sequence 308 AA;

Query Match 81.0%; Score 1286; DB 20; Length 308;

Best Local Similarity 80.8%; Pred. No. 3.8e-126;

Matches 248; Conservative 24; Mismatches 31; Indels 4; Gaps 2;

QY 4 APRVISDKASMRSWRSMEAOQKLGIVPTMGFLHAGHLSLYAQAQRLSDVAVSIYVNP 63

DB 2 APWISDKDEMRSKRSMSQKLGIVPTMGFLHAGHLSLYAQAQRLSDVAVSIYVNP 61

QY 64 GQFAPTEDLSTYPSDFDGDVKKLASVPGGVGVVFFHPRNLYDYKNGGGDVABAG--GMVS 121

DB 62 GQFSPTEDLSTYPSDFDGDVKKLASVPGGVGVVFFHPRNLYDYKNGGGDVABAGCGGGVVS 121

QY 122 CVE--SGSHESWVRVEKLELGLCGKSRPVFRGVATVTKLFNIVEPDVAVFGKKDYQ 179

DB 122 CVDRRSGFGHEWVRAEKLKLEKLPKSRPVFRGVATVTKLFNIVEPDVAVFGKKDYQ 181

QY 180 WELIQRWVRLDLSFKVIGAEITRDNDGLAMSSRNVLSPEREKALSINKSLRAKSA 239

DB 182 WKLIQRWVRLDLSFKVIGAEITRDNDGLAMSSRNVLSPEREKALSINKSLRAKSA 241

QY 240 GDGQVHCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPWFCVAAWFGKVR 299

DB 242 EDGQIHCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPWFCVAAWFGKVR 299

QY 300 DNMEINL 306

DB 302 DNIEINL 308

RESULT 2

ABB93747

ID ABB93747 standard; Protein; 310 AA.

AC ABB93747;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2958.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB ) BAYER AG.

XX PT Tietjen K, Weidler M;

XX DR WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds,

XX PT comprising aligning and comparing nucleic acid or amino acid sequences

XX PT from plant with nucleic acid or amino acid sequences from non-plant

XX PT organisms -

XX PS Claim 5; SEQ ID NO 2958; 261pp + Sequence Listing; English.

XX CC The invention relates to identifying target proteins

XX CC (AB90790-AB94016) for herbicidally active compounds, comprising

XX CC aligning and comparing nucleic acid or amino acid sequences from plant

XX CC with nucleic acid or amino acid sequences from non-plant organisms using

XX CC suitable search parameters, where plant sequences having an E-value

XX CC greater by a factor of 3 than the E-value of most similar non-plant

XX CC sequences are selected. The polypeptides or nucleic acids encoding them

XX CC are useful for identifying modulators. The identified modulators are

XX CC useful as herbicides.

XX SQ Sequence 310 AA;

Query Match 69.3%; Score 1101; DB 23; Length 310;

Best Local Similarity 67.2%; Pred. No. 1e-106;

Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;

QY 5 PRVISDKASMRSWRSMEAOQKLGIVPTMGFLHAGHLSLYAQAQRLSDVAVSIYVNP 64

DB 3 PEVIRDKDSMRKNSRMSQKLTGLVPTMGFLHAGHLSLYAQAQRLSDVAVSIYVNP 62

QY 65 QFAPTEDLSTYPSDFDGDVKKLASVPGGVGVVFFHPRNLYDYKNGGGDVABAG 118

DB 63 QFSPTEDLSTYPSDFDGDVKKLASVPGGVGVVFFHPRNLYDYKNGGGDVABAG 119

QY 119 MVSQVSSG--CHESWVRVEKLELGLCGKSRPVFRGVATVTKLFNIVEPDVAVFGKKDY 177

DB 120 VVSVEEGGLGHETWIRVERLEKFCGKSRPVFRGVATVTKLFNIVEPDVAVFGKKDY 179

QY 178 QWELIQRWVRLDLSFKVIGAEITRDNDGLAMSSRNVLSPEREKALSINKSLRAKSA 237

DB 180 QWELIQRWVRLDLSFKVIGAEITRDNDGLAMSSRNVLSPEREKALSINKSLRAKSA 239

QY 238 AAGDQVHCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPWFCVAAWFGKVR 297

DB 240 SVAEGKTNCARLKDMLIQVVGSGRVDYVEIVDQETLEGVEIKSGVVICVAAWFGTVR 299

QY 298 LIDNMEINLSM 308

DB 300 LIDNIEINLSL 310

RESULT 3

AAG39461

ID AAG39461 standard; Protein; 643 AA.

AC AAG39461;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 48826.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX

PR 25-FEB-2000; 200EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130445.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139494.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142330.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145324.  
PR 26-JUL-1999; 99US-0145326.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148241.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 27-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.

PR	28-SEP-1999;	99US-0156458.	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 48825.
PR	29-SEP-1999;	99US-0156596.	DE	
PR	04-OCT-1999;	99US-0157117.	XX	
PR	05-OCT-1999;	99US-0157753.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	06-OCT-1999;	99US-0157865.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	07-OCT-1999;	99US-0158029.	KW	termination sequence.
PR	08-OCT-1999;	99US-0158232.	XX	
PR	12-OCT-1999;	99US-0158369.	OS	Arabidopsis thaliana.
PR	13-OCT-1999;	99US-0159293.	XX	
PR	13-OCT-1999;	99US-0159294.	XX	
PR	13-OCT-1999;	99US-0159295.	XX	
PR	14-OCT-1999;	99US-0159329.	XX	
PR	14-OCT-1999;	99US-0159330.	XX	
PR	14-OCT-1999;	99US-0159331.	XX	
PR	14-OCT-1999;	99US-0159637.	XX	
PR	14-OCT-1999;	99US-0159638.	XX	
PR	18-OCT-1999;	99US-0159584.	PR	25-FEB-1999;
PR	21-OCT-1999;	99US-0160741.	PR	99US-0121825.
PR	21-OCT-1999;	99US-0160767.	PR	99US-0123180.
PR	21-OCT-1999;	99US-0160768.	PR	09-MAR-1999;
PR	21-OCT-1999;	99US-0160770.	PR	99US-0123548.
PR	21-OCT-1999;	99US-0160814.	PR	99US-0125788.
PR	21-OCT-1999;	99US-0160815.	PR	25-MAR-1999;
PR	22-OCT-1999;	99US-0160980.	PR	99US-0126264.
PR	22-OCT-1999;	99US-0160981.	PR	29-MAR-1999;
PR	22-OCT-1999;	99US-0160989.	PR	99US-0126785.
PR	25-OCT-1999;	99US-0161404.	PR	01-APR-1999;
PR	25-OCT-1999;	99US-0161405.	PR	99US-0127462.
PR	25-OCT-1999;	99US-0161406.	PR	06-APR-1999;
PR	26-OCT-1999;	99US-0161359.	PR	99US-0128234.
PR	26-OCT-1999;	99US-0161360.	PR	99US-0128714.
PR	26-OCT-1999;	99US-0161361.	PR	99US-0128845.
PR	28-OCT-1999;	99US-0161920.	PR	99US-0130077.
PR	28-OCT-1999;	99US-0161920.	PR	99US-0130077.
PR	28-OCT-1999;	99US-0161992.	PR	99US-0130449.
PR	28-OCT-1999;	99US-0161993.	PR	99US-0130510.
PR	29-OCT-1999;	99US-0162142.	PR	99US-0130891.
PR			PR	99US-0131449.
PR			PR	99US-0132048.
PR			PR	99US-0132407.
PR			PR	99US-0132407.
PR			PR	99US-0132484.
PR			PR	99US-0132485.
PR			PR	99US-0132486.
PR			PR	99US-0132487.
PR			PR	99US-0132863.
PR			PR	99US-0134256.
PR			PR	99US-0134256.
PR			PR	99US-0134218.
PR			PR	99US-0134219.
PR			PR	99US-0134221.
PR			PR	99US-0134370.
PR			PR	99US-0134768.
PR			PR	99US-0134941.
PR			PR	99US-0135124.
PR			PR	99US-0135353.
PR			PR	99US-0135629.
PR			PR	99US-0136021.
PR			PR	99US-0136392.
PR			PR	99US-0136782.
PR			PR	99US-0137222.
PR			PR	99US-0137528.
PR			PR	99US-0137502.
PR			PR	99US-0137724.
PR			PR	99US-0138094.
PR			PR	99US-0138540.
PR			PR	99US-0138847.
PR			PR	99US-0139119.
PR			PR	99US-0139452.
PR			PR	99US-0139453.
PR			PR	99US-0139492.
PR			PR	99US-0139454.
PR			PR	99US-0139455.
PR			PR	99US-0139456.
PR			PR	99US-0139457.
PR			PR	99US-0139458.
PR			PR	99US-0139459.
PR			PR	99US-0139460.
PR			PR	99US-0139461.
PR			PR	99US-0139462.
PR			PR	99US-0139463.
PR			PR	99US-0139750.
PR			PR	99US-0139763.

Query Match

Best Local Similarity

Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;

5 PRVISDKASMRSWRSRAQKLGIVPTMGFLHAGLSLVAQARQLSDVVASIYVNP 64

336 PEVTRDKSMKRWKRAMRSQKKTGLVPTMGILHAGLSLVRQSLALDVTVTWVSIYVNP 395

65 QFAPTEDLSTYPSDFDGVKKLASVPGGVVVHPRNLYDYG-----KNGGGDVARAGG 118

396 QFSPTEDLSTYPSDFSGDLTKLAALSGKGVVVFNPNLYDYGGETKKINDGGN--GGR 452

119 MVSCVESGS-CHESWVRVEKLELGLCGKSRPVFRGVAIVTKLFNIVEPDVAVFGKDY 177

453 VVSCVEEGLGHETWIRVERLEKFCGKSRPVFRGVAIVTKLFNIVEPDVAVFGKDY 512

178 QQWELIQWRVRLDPSIKVGAETITRDNDGLAMSSRNVHLSPERERKALSINKSLRAKS 237

513 QQWELIQWRVRLDPSIKVGAETITRDNDGLAMSSRNVHLSPERERKALSINKSLRAKS 572

238 AAGGQVHCETKLNLIQSVTDAGRIDYABIVDQNNLEKVEIKSPVVFCAAWFGKVR 297

573 SVARGKTNCAELKDMIIQQVVGSGRVVDYVIVDQETLEGVEIKSGWVICVAWFGVVR 632

298 LIDNNEINLSM 308

633 LIDNNEINLSM 643

RESULT 4

AAG39460

ID AAG39460 standard; Protein; 653 AA.

XX

AC AAG39460;

XX

DT 18-OCT-2000 (first entry)

```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 25-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 27-JUN-1999; 99US-0140991.
PR 28-JUN-1999; 99US-0141287.
PR 29-JUN-1999; 99US-0141842.
PR 30-JUN-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 03-JUL-1999; 99US-0142803.
PR 04-JUL-1999; 99US-0142920.
PR 05-JUL-1999; 99US-0142977.
PR 06-JUL-1999; 99US-0143542.
PR 07-JUL-1999; 99US-0143624.
PR 08-JUL-1999; 99US-0144005.
PR 09-JUL-1999; 99US-0144085.
PR 10-JUL-1999; 99US-0144086.
PR 11-JUL-1999; 99US-0144325.
PR 12-JUL-1999; 99US-0144331.
PR 13-JUL-1999; 99US-0144332.
PR 14-JUL-1999; 99US-0144333.
PR 15-JUL-1999; 99US-0144334.
PR 16-JUL-1999; 99US-0144335.
PR 17-JUL-1999; 99US-0144632.
PR 18-JUL-1999; 99US-0144884.
PR 19-JUL-1999; 99US-0144814.
PR 20-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145089.
PR 24-JUL-1999; 99US-0145192.
PR 25-JUL-1999; 99US-0145145.
PR 26-JUL-1999; 99US-0145218.
PR 27-JUL-1999; 99US-0145224.
PR 28-JUL-1999; 99US-0145276.
PR 29-JUL-1999; 99US-0145913.
PR 30-JUL-1999; 99US-0145918.
PR 31-JUL-1999; 99US-0145919.
PR 01-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 03-AUG-1999; 99US-0146388.
PR 04-AUG-1999; 99US-0146389.
PR 05-AUG-1999; 99US-0147038.
PR 06-AUG-1999; 99US-0147204.
PR 07-AUG-1999; 99US-0147302.
PR 08-AUG-1999; 99US-0147192.
PR 09-AUG-1999; 99US-0147260.
PR 10-AUG-1999; 99US-0147303.
PR 11-AUG-1999; 99US-0147416.
PR 12-AUG-1999; 99US-0147433.
PR 13-AUG-1999; 99US-0147935.
PR 14-AUG-1999; 99US-0148171.
PR 15-AUG-1999; 99US-0148319.
PR 16-AUG-1999; 99US-0148565.
PR 17-AUG-1999; 99US-0148341.
PR 18-AUG-1999; 99US-0149426.
PR 19-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 21-AUG-1999; 99US-0149929.
PR 22-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 24-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 28-AUG-1999; 99US-0151303.
PR 29-AUG-1999; 99US-0151438.
PR 30-AUG-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 02-SEP-1999; 99US-0153070.
PR 03-SEP-1999; 99US-0153758.
PR 04-SEP-1999; 99US-0154018.
PR 05-SEP-1999; 99US-0154039.
PR 06-SEP-1999; 99US-0154779.
PR 07-SEP-1999; 99US-0155119.
PR 08-SEP-1999; 99US-0155486.
PR 09-SEP-1999; 99US-0155659.
PR 10-SEP-1999; 99US-0156458.
PR 11-SEP-1999; 99US-0156596.
PR 12-SEP-1999; 99US-0157117.
PR 13-SEP-1999; 99US-0157753.
PR 14-SEP-1999; 99US-0157865.
PR 15-SEP-1999; 99US-0158029.
PR 16-SEP-1999; 99US-0158232.
PR 17-SEP-1999; 99US-0158369.
PR 18-SEP-1999; 99US-0159293.
PR 19-SEP-1999; 99US-0159294.
PR 20-SEP-1999; 99US-0159295.
PR 21-SEP-1999; 99US-0159329.
PR 22-SEP-1999; 99US-0159330.
PR 23-SEP-1999; 99US-0159331.
PR 24-SEP-1999; 99US-0159637.
PR 25-SEP-1999; 99US-0159638.
PR 26-SEP-1999; 99US-0159584.
PR 27-SEP-1999; 99US-0160741.
PR 28-SEP-1999; 99US-0160767.
PR 29-SEP-1999; 99US-0160768.
PR 30-SEP-1999; 99US-0160770.
PR 01-OCT-1999; 99US-0160814.
PR 02-OCT-1999; 99US-0160815.
PR 03-OCT-1999; 99US-0160980.
PR 04-OCT-1999; 99US-0160981.
PR 05-OCT-1999; 99US-0160989.
PR 06-OCT-1999; 99US-0161404.
PR 07-OCT-1999; 99US-0161405.
PR 08-OCT-1999; 99US-0161406.
PR 09-OCT-1999; 99US-0161359.
PR 10-OCT-1999; 99US-0161360.
PR 11-OCT-1999; 99US-0161361.
PR 12-OCT-1999; 99US-0161320.
PR 13-OCT-1999; 99US-0161592.
PR 14-OCT-1999; 99US-0161993.
PR 15-OCT-1999; 99US-0162142.

Query Match 69.3%; Score 1101; DB 21; Length 653;
Best Local Similarity 67.2%; Pred. No. 3.2e-106;
Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;

QY 5 PRVISDKMRSRNRRAQKGLIGLPTMGFLHAGHLSLVAQARQLSDVVAVSIYVPG 64
Db 346 PEVIRKDSMRKWSRAMRSQKGTIGLPTMGFLHAGHLSLVRQSLATDVTVSIYVPG 405
QY 65 QFAPTEDLSTYPSDFDGVKKLASVPGGVDDVPHRNLYDYG-----KNGGGDVABAGG 118
Db 406 QFSPTEDLSTYPSDFSGDLTKLAALSGKGVVFNPKNLYDYGGETKKINDGGCN---GGR 462
QY 119 MVSCVBSGS-GHESVVRVEKLEGLCGKSRPVPFRGVATVTKLFNIVEPDVAVFGKKDY 177
Db 463 VVSCVEGGHGHETWIRVERLEKFGCKSRPVPFRGVATVTKLFNIVEPDVAVFGKKDY 522
QY 178 QQWRLIQRMVRDLDFSIKIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLLRKS 237
Db 523 QQWRLIQRMVRDLNFGIEIVGSDIAREKDGGLAMSSRNHLSDEERORALSISRLAMAKA 582
QY 238 AAGDQVHCHEKLTNLVIQSVTDAGGRIDYARIVDQNNLEKVEQIKSPVVFCAVAFGKVR 297
Db 583 SVAEGKTNCAELKDMIIQQVWGSAGRDYVEIVDQETLEGVEEIKSGVVICVAAWFGTVR 642
```

QY 298 LIDNMEINLSM 308  
| | | | | :  
Db 643 LIDNIEINVS 653

RESULT 5  
AAG39459  
ID AAG39459 standard; Protein; 654 AA.  
XX AC AAG39459;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 48824.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.





PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.

```
PR 28-OCT-1999; 99US-01611993.
PR 29-OCT-1999; 99US-0162142.

Query Match          69.1%; Score 1098; DB 21; Length 310;
Best Local Similarity 67.2%; Pred. No. 2.1e-106;
Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;

Qy 5 PRVSDKASMRWSRRAQKLGIVPTMGFLHAGHLSLVAQARQLSDVVAIVYVNP 64
Db 3 PEVIRDKDSMRKWRARSRQKTLGLVPTMGFLHAGHLSLVRQSLALTDVTVWSIYVNP 62

Qy 65 QFATEDLSTVPSDFDGVKKLASVPGVDVVFHPRNLYDYG-----KGGGDVAEAG 118
Db 63 QFSTEDLSTVPSDFDGLTKLAALSGKVVVFHPRNLYDYGTTKINDGGN---GGR 119

Qy 119 MVSCVESGS-GHESWVRVEKLELGCGKSRVFRFGVATVTKLFNIVEPDAVFGKKDY 177
Db 120 VVSCVEEGGLHETWIRVERLEKLGCGKSRVFRFGVATVTKLFNIVEPDAVFGKKDY 179

Qy 178 QOWRLIQRMVRDLDFSIKIVGAETTRNDGLAMSSRVNHLSPEREKALSINKSLRAKS 237
Db 180 QOWRIIQRMVRDLNFGIEIVGSDIAREKDGAMSSRVNHLSDERQALSISRSLAMAKA 239

Qy 238 AAGDGVHCEKLTNLVIQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVR 297
Db 240 SVAEGKNCBAELKDMIIQQVVGSRVDYVEIVDQETLEGVEEIKSGVVICVAAWFGTVR 299

Qy 298 LIDNMEINLSM 308
Db 300 LIDNIEINSL 310

RESULT 7
AAG28280
ID AAG28280 standard; Protein; 322 AA.
XX
AC AAG28280;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33439.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
```

```
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 69.1%; Score 1098; DB 21; Length 322;
Best Local Similarity 67.2%; Pred. No. 2.2e-106;
Matches 209; Conservative 46; Mismatches 46; Indels 10; Gaps 3;

QY 5 PRVISDKASMRSWSRMRAQGLIGLVPTMGFLHAGHLSLVAQARQLSDVAVSIYVNP 64
Db 15 PEVIRKDSMEKWSRAMRSQKTIQGLVPTMGYLGHLHLVRSQSLATDVTWVSIYVNP 74
QY 65 QFAPTEDLSTYPSDFDGDVKKLASVPGGVVPHRNLYDYG-----KNGGGDVAEAGG 118
Db 75 QFSPTEDLSTYPSDFSGDLTKLAALSGGKVVVFNPKNLYDYGETKKINDGGN---GGR 131
QY 119 MVSCVESGS-CHESWVRVEKLELGLCGKSRPVFRGVATVTKLFNIVEPDPVAFGKKDY 177
Db 132 VVSCVEGGGLGHETWIRVERLEKXLCGKSRPVFRGVATVTKLFNIVEPDPVAFGKKDY 191
QY 178 QQWRLIQRMVRDLDFSIKVIGABEITRDNDGLAMSSRNVLHSPEREKALSINKSLRAKS 237
Db 192 QQWRLIQRMVRDLNFGIEIVGSDIAREKDGAMSSRNVLHSPEREKALSINKSLRAK 251
QY 238 AAGDQVHCHEKLTNLVTQSVTDAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAAFQKVR 297
Db 252 SVAEGKTNCABLKDMMIIQQVVGSGRVDYVEIVDQETLEGVEEIKSGVVICVAAAFQKVR 311
QY 298 LIDNWEINLSM 308
Db 312 LIDNIEINVS 322

RESULT 8
AAG28282
ID AAG28282 standard; Protein; 299 AA.
XX AC AAG28282;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33441.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
```

PR 01-APR-1999; 99US-0137462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 21-MAY-1999; 99US-0135623.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140334.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142303.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149302.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 26-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.

```

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 14-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 67.4%; Score 1071; DB 21; Length 299;
Best Local Similarity 67.2%; Pred. No. 1.4e-103;
Matches 203; Conservative 46; Mismatches 43; Indels 10; Gaps 3;

QY 14 MRSWSRMRRAQKLGVLPTMGFLHAGHLSLVAAQRLSDVAVSIYVNPQGFAPTEDL 73
Dd 1 MEKWSRMRSQKTLGLVPTMGVLYHGHLSLVRSQSLATDVTVSIYVNPQGFAPTEDL 60

QY 74 TYPSPDFDGVKKLASVPGGVVVFHPRNLVDYV-----KNGGGDVAAAGGMVSCVESGS 127
Dd 61 TYPSPDFSGDLTKLAALSGGVVVFHPRNLVDYVGTETTKINDGGN---GGRVWVSCVEGG 117

QY 128 -GHESWVVEKLELGLCGKSRPFRGVATVTKLFNIVEPDVAVFGKDYQWRLIQRM 186
Dd 118 LGHETWIRVERLEKXKLGKSRPFRGVATVTKLFNIVEPDVAVFGKDYQWRLIQRM 177

QY 187 VRDLFSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKGAAGGVHVC 246
Dd 178 VRDLNFGIEIVGSDIAREKDGGLAMSSRNHLSPEEREKALSINKSLRAKGAAGGVHVC 237

QY 247 EKLTLNLVQSVTDAGRIDYAEIVDQNNLEKVEIKSPVFCVAAWFGKVRLLDNMEINL 306
Dd 238 AELKDMIIQQVVGSGAGRDVVEIVDQETLEGVEEIKSGVVICVAAWFGVTVRLIDNIEIN 297

QY 307 SM 308
Dd 298 SL 299

RESULT 9
AA42403
ID AA42403 standard; Protein; 333 AA.
XX
AC AA42403;
XX
DT 02-DEC-1999 (first entry)
XX
DE Amino acid sequence of Rice Pantothenate Synthetase (PS).
XX
KW pantothenate synthetase; cloning; herbicide; biosynthesis;
XX
KW Coenzyme A precursor; pantoate-beta-alanine ligase.
XX
OS Oryza sativa.
XX

```

```

PN WO9942565-A1.
XX
PD 26-AUG-1999.
XX
PF 02-JUN-1998; 98WO-EP03261.
XX
PR 31-MAY-1997; 97GB-0011163.
PR 27-JUN-1997; 97GB-0013477.
XX
PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.
XX
XX
PI Abell C, Smith AG, Genschel U, Laber B;
DR WPI; 1999-527466/44.
DR N-PSDB; AAZ20927.
XX
PT New isolated DNA molecule, useful in herbicidal compositions -
XX
PS Claim 6; Fig 2.2; 64pp; English.
XX
CC This is the amino acid sequence of the Oryza sativa (rice) Pantothenate
CC synthetase (PS) enzyme, which is essential in the biosynthesis of
CC Coenzyme A.
CC The invention provides methods to assay proteins for PS activity, and
CC the ability of compounds to inhibit PS.
CC The assay compounds which are found to inhibit PS can be used as
CC herbicides.
CC The nucleotide probe can be used to amplify PS coding sequences from a
CC chosen organism via the process of polymerase chain reaction.
CC A number of assays have previously been used to measure PS activity
CC but are unsuitable for large screen biochemical screening of compounds
CC to find PS inhibitors. In this invention however the isolation of the
CC DNA molecule encoding PS enables a host cell to produce the protein
CC which can then be tested against compounds potentially able to inhibit
CC the enzyme.
XX
SQ Sequence 333 AA;

Query Match 58.4%; Score 927.5; DB 20; Length 333;
Best Local Similarity 59.9%; Pred. No. 1.9e-88;
Matches 185; Conservative 51; Mismatches 60; Indels 13; Gaps 6;

QY 5 PRVTSKASMRSWRSRRAQKLGVLPTMGFLHAGHLSLV--AAQRLSDVVA--VSIY 60
Dd 27 PEVIRDKAAWRSRRRAEGKTAVVPTMGVLYHGHLSLVASAAAASADPVAIVVTIY 86

QY 61 VNPQGFAPTEDLSTYPSDFDGVKKLASVPGGVVVFHPRNLX--DYKNGGGDVAAAGG 118
Dd 87 VNPQGFAPSEDLATYPSDFAGDLKLAST-GVVDVAFNPFDLYVRGAGRGAG---SGG 141

QY 119 MVSCVE--SGSGHESWVVEKLELGLCGKSRPFRGVATVTKLFNIVEPDVAVFGKDD 176
Dd 142 AISCLEAAGDGCHETWVRVERLEKGLCGASRPFRGVATVSKLFNIIIPDPVFOKDD 201

QY 177 YQWRLIQRMVVRDLDFSIVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAK 236
Dd 202 YQWVRVILPYWGLDFGIEIMGSRNCARTDGLAMSSRNHLSPEEREKALSINKSLRAK 261

QY 237 SAAGDGQVHCEKLTNLVSIQSVTDAGRIDYAEIVDQNNLEKVEIKSPVFCVAAWFGKV 296
Dd 262 TGAUKGNNTSKQIKNKIVQTLTETGGQVDYVEIVEQESLVPVEQIDGVPVVCVAAWFGKV 321

QY 297 RLIDNMEIN 305
Dd 322 RLIDNIEID 330

RESULT 10
AAU01245
ID AAU01245 standard; Protein; 286 AA.
XX
XX AAU01245;
XX

```

DT 18-JUL-2001 (first entry)  
XX B. subtilis pantothenate synthetase.  
XX  
XX Pantothenate synthetase; panC; pantothenate biosynthesis;  
KW vitamin B5; nutritional supplement; panto-compound; pantoate.  
XX Bacillus subtilis.  
OS WO200121772-A2.  
XX PN 29-MAR-2001.  
XX PD  
XX XX  
PF 21-SEP-2000; 200WO-US25993.  
XX PR 21-SEP-1999; 99US-0400494.  
PR 07-JUN-2000; 2000US-0210072.  
PR 28-JUL-2000; 2000US-0221836.  
PR 24-AUG-2000; 2000US-0227860.  
XX PA (OMNI-) OMNIGENE BIOPRODUCTS.  
XX PI Yocum RR, Patterson TA, Hermann T, Pero JG;  
XX WPI; 2001-218644/22.  
DR N-PSDB; AAS00989, AAS02314.  
XX  
PT New recombinant microorganism which overexpress a Bacillus subtilis  
PT pantothenate biosynthetic enzyme, useful for the high yield production  
PT of panto-compounds such as pantothenate and pantoate -  
XX Example 1; Page 172-173; 292pp; English.  
XX The sequence is B. subtilis pantothenate synthetase, encoded by the  
CC panC gene, an enzyme of the pantothenate biosynthetic pathway.  
CC Pantothenate, also known as vitamin B5, is used as a nutritional  
CC supplement in mammals and humans. The invention concerns methods of  
CC producing recombinant microorganisms overexpressing at least one Bacillus  
CC subtilis pantothenate biosynthetic enzyme. The microorganisms and methods  
CC of producing them are useful for producing a panto-compound such as  
CC pantothenate or pantoate, which is a nutritional requirement for  
CC livestock and humans. The methods are also useful for the identification  
CC of pantothenate kinase modulators. Panto-compounds are produced at a  
CC significantly higher yield than prior art methods and can be produced  
CC independent of the need to feed precursors which decreases expense.

SQ Sequence 286 AA;

Query Match 31.1%; Score 494; DB 22; Length 286;  
Best Local Similarity 36.6%; Pred. No. 5e-43;  
Matches 111; Conservative 51; Mismatches 115; Indels 26; Gaps 3

QY 6 RVISDKASMRSWSMRAQGKLGIVPTMGFLHGHLSLVAQAQLSDVAVSIYNPGQ 65  
Db |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
2 RQUTDISOLKEAIKOYHSEGSIGFVPVTMGFLHEGHLTLADKARQENDAVTMSIFWNPQA 61  
QY 66 FAPTEDLSTVPSDFDGDVKKLASVPGGVDDVVYP--RNLYDYGNKGSGDVAEAGMWSCV 123  
Db |||||:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
62 FGNEDEPEAFPRDLERDAALAE--AGVDILFTPDADMVGPGEKN----- 104

QY 124 ESGSGHSWVRVEKLEGLCGSKSRPFRGRVATVTKLFNTIPEDPVAVFGKKDYCOWRLI 183  
Db :||:||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
105 -----VTIHVERRTDVLGORSEGHFDGVAIVLTCLFNLVKTRAYFLGKDAAQQVAV 157

QY 184 QRMVRDDLFSIKVIABITRDNCDGLAMSSRNVLHSPEREKALSINKSLLRSAAGDGQ 243  
Db ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
158 DGLISDFPMDELFPVDTVREEDGLAKSSRNVLTAERKEAPKLYRALQTSABLVOAGE 217

QY 244 VHECKLTNLVIQSVDTAGGRIDYAIETVDONNLEKVEQIKSPVFCVAAWFOKRVLIDNM 303  
Db ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
218 RDPEAVIKAANDIIETTSGTIDYVELYSYPELEPNFIAGWMILAFAVFASKRLIDNII 277

QY 304 INL 306



Db 78 LALLEE--AGVDIVFAP-----DVEEMPGLPLV-----WARTSGT 114  
 Qy 141 GLCKSRPFRGVATVTKLFNIVEPDVAVFGKKDYQOMRLIQRMVRDLDFSIKVGAE 200  
 Db 115 KLEGASRPGHFDGVATVAKLFNLRPDRAVFGQKDAQQVAVIRRLVADLDIPVEIRVP 174  
 Qy 201 ITRNDGLAMSRNHLSPEREKALSINKSL--LRAKSAGDGQGVHCEKLTNLIQSVT 258  
 Db 175 IIRGADGLAESRNRQLSADQRAQALVLPQVLSGLQRKAAGEA-----LDIQAR 225  
 Qy 259 DA-----GGRIDYAEIVDQNNLEKVE---QIKSPVVFCAVAFGKVRLLDNMEI 304  
 Db 226 DTLASADGVRDLHLEIVDPATLPLEIDGLLTQPALVVGAI FVGPRVRLIDNIEL 279

## RESULT 14

AAB10034  
 ID AAB10034 standard; Protein; 279 AA.

AC AAB10034;

DT 02-NOV-2000 (first entry)

DE C. glutamicum panC protein.

XX D-pantothenic acid; panB; panC; ilvD; pantotheanate synthetase;  
 KW ketopantoathydroxymethyltransferase; dihydroxyaciddehydratase;  
 KW panBC operon; vitamin.

XX Corynebacterium glutamicum.

XX EP1006189-A2.

XX 07-JUN-2000.

XX 30-NOV-1999; 99EP-0123738.

XX 01-DEC-1998; 98DE-1055312.

XX (DEGS ) DEGUSSA-HUELS AG.

PA (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Eggeling L, Thierbach G, Sahn H;

XX WPI; 2000-378263/33.

DR N-PSDB; AAA40282.

PT Recombinant Corynebacterium DNA useful for production of pantothenic  
 PT acid vitamin, comprises panB, panC or ilvD genes encoding enzymes -

XX Example 1; Page 16-17; 27pp; German.

XX This invention describes novel recombinant Corynebacterium DNA (I),  
 CC present in microorganisms of the Corynebacterium genus and comprising  
 CC at least one of the panB (ketopantohydroxymethyltransferase), panC  
 CC (pantothenic acid synthetase), especially the panBC operon, and/or ilvD  
 CC (dihydroxyaciddehydratase) genes. (I) is useful for the preparation of  
 CC pantothenic acid a vitamin which has applications including cosmetics,  
 CC medicine and human and animal nutrition. The new preparation method using  
 CC fermentation techniques produces the required stereo-isomer D form of  
 CC pantothenic acid. This sequence represents the Corynebacterium glutamicum  
 CC panC protein which is described in the method of the invention.

XX Sequence 279 AA;

Query Match 28.7%; Score 455.5; DB 21; Length 279;

Best Local Similarity 41.2%; Pred. No. 5.3e-39;

Matches 121; Conservative 33; Mismatches 93; Indels 47; Gaps 9;

Qy 26 KLIGLVPTMGTHAGHLSLVAAQRLSDVAVSIYVNPQF---APTEDLSTYSPDPPGD 82

Db 18 KSVGLVPTMGALHSGHSLVAAKRAENDTVVASIFVNPQLQFEALGDCDDYRNPRLDAD 77

Qy 83 VKKLASVPGVDVVVHFHPRNLYDYKNGGGDVAE--AGMVSVCVBSGSHSVRVEKLEL 140  
 Db 78 LALLEE--AGVDIVFAP-----DVEEMPGLPLV-----WARTSGT 114  
 Qy 141 GLCKSRPFRGVATVTKLFNIVEPDVAVFGKKDYQOMRLIQRMVRDLDFSIKVGAE 200  
 Db 115 KLEGASRPGHFDGVATVAKLFNLRPDRAVFGQKDAQQVAVIRRLVADLDIPVEIRVP 174  
 Qy 201 ITRNDGLAMSRNHLSPEREKALSINKSL--LRAKSAGDGQGVHCEKLTNLIQSVT 258  
 Db 175 IIRGADGLAESRNRQLSADQRAQALVLPQVLSGLQRKAAGEA-----LDIQAR 225  
 Qy 259 DA-----GGRIDYAEIVDQNNLEKVE---QIKSPVVFCAVAFGKVRLLDNMEI 304  
 Db 226 DTLASADGVRDLHLEIVDPATLPLEIDGLLTQPALVVGAI FVGPRVRLIDNIEL 279

## RESULT 15

AAG89878

ID AAG89878 standard; Protein; 279 AA.

XX AC AAG89878;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 3632.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

DR N-PSDB; AAH65097.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analysing  
 PT expression profile or pattern of a gene and identifying homologous gene

XX Claim 17; SEQ ID NO: 3632; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein  
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of coryneform bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Coryneform bacterium, and identifying a homologue of a gene derived  
 CC from coryneform bacterium. Coryneform bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described  
 CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.

XX Sequence 279 AA;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:34:50 ; Search time 31 Seconds  
(without alignments)  
1859.837 Million cell updates/sec

Title: US-10-033-269-9  
Perfect score: 1588  
Sequence: 1 MAPAPRVISDKASMRWSRS.....AWFGKRLIDNMEINLSMNV 310

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 18593659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1588	100.0	310	12	US-10-033-269-9
2	1286	81.0	308	12	US-10-033-269-15
3	1164	73.3	311	12	US-10-033-269-7
4	1002	63.1	316	12	US-10-033-269-13
5	961.5	60.5	325	12	US-10-033-269-2
6	959.5	60.4	323	12	US-10-033-269-4
7	927.5	58.4	313	12	US-10-033-269-14
8	911.5	57.4	296	12	US-10-033-269-11
9	455.5	28.7	279	10	US-09-738-626-3632
10	396	24.9	333	15	US-10-156-761-12221
11	231	14.5	77	15	US-10-113-948-2
12	228	14.4	266	10	US-09-738-626-6455
13	178	11.2	79	15	US-10-113-948-1
14	94	5.9	659	15	US-10-128-714-8189
15	93.5	5.9	423	15	US-10-127-032-140

16	91	5.7	339	10	US-09-738-626-4524	Sequence 4524, Ap
17	89	5.6	801	15	US-10-156-761-12150	Sequence 12150, A
18	89	5.6	1387	15	US-10-156-761-13000	Sequence 13000, A
19	85.5	5.4	224	10	US-09-993-170-10	Sequence 10, Appl
20	85.5	5.4	385	12	US-10-043-639A-8	Sequence 8, Appl
21	85.5	5.4	630	12	US-09-882-227-450	Sequence 460, App
22	85.5	5.4	715	15	US-10-156-761-9360	Sequence 9360, Ap
23	85	5.4	479	11	US-09-934-455-210	Sequence 210, App
24	84.5	5.3	223	10	US-09-993-170-3	Sequence 3, Appl
25	84.5	5.3	224	10	US-09-993-170-13	Sequence 13, Appl
26	84.5	5.3	565	14	US-10-062-254-92	Sequence 92, Appl
27	84	5.3	706	15	US-10-270-333-133	Sequence 153, App
28	83.5	5.3	2756	15	US-10-331-061-7	Sequence 7, Appl
29	82.5	5.2	5147	12	US-10-174-677-4	Sequence 4, Appl
30	82	5.2	583	10	US-09-758-269-2	Sequence 2, Appl
31	82	5.2	1090	9	US-09-833-435A-5	Sequence 5, Appl
32	82	5.2	1090	12	US-10-375-720-5	Sequence 5, Appl
33	82	5.2	1091	11	US-09-262-126C-4	Sequence 4, Appl
34	82	5.2	1091	15	US-10-245-803-4	Sequence 4, Appl
35	82	5.2	5183	12	US-10-107-521-1	Sequence 1, Appl
36	81.5	5.1	223	10	US-09-993-170-7	Sequence 7, Appl
37	81.5	5.1	361	12	US-10-166-225A-65	Sequence 65, Appl
38	81.5	5.1	691	12	US-10-321-802-22	Sequence 22, Appl
39	80.5	5.1	403	10	US-09-864-921-176	Sequence 176, App
40	80.5	5.1	900	9	US-09-815-242-13381	Sequence 13381, A
41	80.5	5.1	900	9	US-09-815-242-13663	Sequence 13663, A
42	80.5	5.1	1009	10	US-09-864-921-107	Sequence 107, App
43	80.5	5.1	4349	12	US-10-174-677-76	Sequence 76, Appl
44	80.5	5.1	4349	12	US-09-970-944-4	Sequence 4, Appl
45	80.5	5.1	4349	12	US-09-970-944-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-10-033-269-9  
; Sequence 9, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase  
; FILE REFERENCE: BB1446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-033-269-9

Query Match	100.0%;	Score 1588;	DB 12;	Length 310;
Best Local Similarity	100.0%;	Pred. No. 1.4e-158;		
Matches 310;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAPAPRVISDKASMRWSRMRQAQKLI	GLVPTMGFLHAGHLSLVAQARQLSDVAVSIY	60
Db	1	MAPAPRVISDKASMRWSRMRQAQKLI	GLVPTMGFLHAGHLSLVAQARQLSDVAVSIY	60
Qy	61	VNPGQFAPTEDISTYPSDFDGVKULASVPGGVDVVFHPRNLYDYKNGGGDVAEAGMV	120	
Db	61	VNPGQFAPTEDISTYPSDFDGVKULASVPGGVDVVFHPRNLYDYKNGGGDVAEAGMV	120	
Qy	121	SCVESGSHSVRRVEKLEGLCGKSRPVFFRCVATVTKLFNIVEPDAVFFKKDYQQW	180	
Db	121	SCVESGSHSVRRVEKLEGLCGKSRPVFFRCVATVTKLFNIVEPDAVFFKKDYQQW	180	

QY 181 RLQRMVRDLDFSIKIVIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAG 240  
Db 181 RLQRMVRDLDFSIKIVIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAG 240  
QY 241 DQGVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVLID 300  
Db 241 DQGVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVLID 300  
QY 301 NMEINLSMNV 310  
Db 301 NMEINLSMNV 310  
RESULT 2  
US-10-033-269-15  
; Sequence 15, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase  
; FILE REFERENCE: BB1446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Lotus japonicus  
US-10-033-269-15

Query Match 81.0%; Score 1286; DB 12; Length 308;  
Best Local Similarity 80.8%; Pred. No. 9.1e-127;  
Matches 248; Conservative 24; Mismatches 31; Indels 4; Gaps 2;  
QY 4 APRVSDKASMRWSRMRQAQKLI GLVPTMGFLHAGHLSLVAQARQLSDVVAVSIYVNP 63  
Db 2 APRVSDKASMRWSRMRQAQKLI ALVPTMGFLHAGHLSLVRDAHNDLVAIVSIYVNP 61  
QY 64 GQFAPTELDLSTYPSDFDGVKVLASVPGGVDDVVFHPRNLYDYGNKGGDVAEAG--GMVS 121  
Db 62 GQFSPTELDLSAYPSDFDGLQKLSVPGGVDDVVFHPRNLYDYGGDGDVAEACGGDGVVS 121  
QY 122 CVE--SGSGHESWVRVEKLEGLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKKDYQQ 179  
Db 122 CVDRSGFGHETWVRAEKLEPLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKKDYQQ 181  
QY 180 WRLIQRMVRDLDFSIKIVIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSA 239  
Db 182 WKLIQRMVRDLDFSIKIVIGSEVIREKDGGLAMSRNVHLSPEEREKAVSINKSLFRKSA 241  
QY 240 GGGQVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVL 299  
Db 242 EDGQVHCEKLTNLVIOQSVITEAGGRIDYAEIVDQNNLEKVEIKGVVFCVSAWFGKARLI 301  
QY 300 DNMEINL 306  
Db 302 DNIEINL 308

RESULT 3  
US-10-033-269-7  
; Sequence 7, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase

; FILE REFERENCE: BB1446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-033-269-7

Query Match 73.3%; Score 1164; DB 12; Length 311;  
Best Local Similarity 73.4%; Pred. No. 6.5e-114;  
Matches 224; Conservative 40; Mismatches 37; Indels 4; Gaps 3;  
QY 5 PRVSDKASMRWSRMRQAQKLI GLVPTMGFLHAGHLSLVAQARQLSDVVAVSIYVNP 64  
Db 8 PLVITDKAEMREWSRMRQAQKLI ALVPTMGFLHAGHLSLVRARRADA VVSVVYVNP 67  
QY 65 QFAPTELDLSTYPSDFDGVKVLASVPGGVDDVVFHPRNLYDYGNK--GGDVAEAGGVSC 122  
Db 68 QFAPSELDLSTYPSDFDGLKRAVPGGVDDVVFQNNLYDYQREVSGSVESDNGSVSC 127  
QY 123 V-ESGSGHESWVRVEKLEGLCGKSRPVFRGVATVVTKLFNIVEPDVAVFGKKDYQOWR 181  
Db 128 LEEKGMGHEAWVRVERLEKMGKSRPVFRGVATVVTKLFNIVEPDVAVFGKKDYQOWR 187  
QY 182 LIQRMVRDLDFSIKIVIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAG 241  
Db 188 IIRRLV-NLDFSIQVIGSEVMRDHGLALSRNVHLSPEEREKALSISRSLRAKSAEK 246  
QY 242 GOVHCEKLTNLVIOQSVTDAGGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVLIDN 301  
Db 247 GQVNCQNLKDSVIOAIQIAGKIDYAEIVDQESLEAVEEIRSPVSCVAAWFGKVLIDN 306  
QY 302 MEINL 306  
Db 307 IEINV 311

RESULT 4  
US-10-033-269-13  
; Sequence 13, Application US/10033269  
; Publication No. US20030167503A1  
; GENERAL INFORMATION:  
; APPLICANT: Gutteridge, Steve  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Orozco, Buddy  
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase  
; FILE REFERENCE: BB1446 US NA  
; CURRENT APPLICATION NUMBER: US/10/033,269  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 60/247,938  
; PRIOR FILING DATE: 2000-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-033-269-13

Query Match 63.1%; Score 1002; DB 12; Length 316;  
Best Local Similarity 65.1%; Pred. No. 7.7e-97;  
Matches 203; Conservative 41; Mismatches 60; Indels 8; Gaps 5;  
QY 2 APAPRVISDKASMRWSRMRQAQKLI GLVPTMGFLHAGHLSLVAQARQLSDVVA--VSI 59  
Db 4 AGSEVIRDKAAMRWSRMRQAQKLI ALVPTMGFLHAGHLSLVSAAA VPGVAVVSVI 63  
QY 60 YVNPQGFAPTELDLSTYPSDFDGVKVLASVPGGVDDVVFHPRNLYDYGNKGGDVAE-- 116

```
Db      64 YVNSQFAPTEDLATYPSDLADGRKLAST-GAHAVNPNPDLTHRGAAVSGRRABAPAG 122
Qy      117 GGVSCVSES-GSGHESVVRVEKLELGLCGKSRPVFFRGVATVVTVKLFNIVEPDPVAVFGKK 175
Db      123 AAASCCLEAGDGHETWIRVERLEKGLCGASRPVFFRGVATVVKLFNIVEPDPVAVFGKK 182
Qy      176 DYQWRLIQRMVRDLDSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRA 235
Db      183 DYQWRLICRMVRDLDFAVEITIGAEIVREADGLAMSSRNHLSPEEREKALSIRSLLNA 242
Qy      236 KSAA-GDQVHCEKLTNLVIOQVTDAGRIDYAEIVDQNNLEKVEQIKSPVVFCAAWFG 294
Db      243 RTAALNNSNSASEHIKQIVQTLTEAGRDYVVEIQESLVPVETIDRPPVICVAAWFG 302
Qy      295 KVLIDNNEINL 306
Db      303 KVLIDNIEIH 314
```

```
RESULT 5
US-10-033-269-2
; Sequence 2, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Zea mays
US-10-033-269-2
```

```
Query Match      60.5%; Score 961.5; DB 12; Length 325;
Best Local Similarity 62.0%; Pred. No. 1.5e-92;
Matches 191; Conservative 49; Mismatches 61; Indels 7; Gaps 4;

Qy      5 PRVISDKASMSRSRMRAGKLGIVPTMGFLHAGHLSLVAQARQLSD--VVAVSIYV 61
Db      11 PEVIRDKAAMRAWSRRRRRAEGKAVLVPTMGFLHAGHLSLSAASAGPIAVVSIYV 70
Qy      62 NPGQFAPTEDLSTYPSDFDGVKKLASVPGGVVVFHPRNLYDYKNGGGGVDAEA-GGMV 120
Db      71 NPSQFAPTEDLATYPSDFAGDLRKLAAAT-GVAAVFCPPDLYVRGSADRPASAAGSGAV 129
Qy      121 SCVESGSG--HESWVRVEKLELGLCGKSRPVFFRGVATVVTVKLFNIVEPDPVAVFGKKDYQ 178
Db      130 SCLDAGGHAHETWIRVERLEKGLCGSSRPVFFRGVATVVKLFNIVEPDPVAVFGKKDYQ 189
Qy      179 QWRLIQRMVRDLDSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 238
Db      190 QWRVICRMVRDLDFAIQVSEVREADGLAMSSRNHLSPEEREKALSIRSLLVDARTA 249
Qy      239 AGDQVHCEKLTNLVIOQVTDAGRIDYAEIVDQNNLEKVEQIKSPVVFCAAWFGKVL 298
Db      250 ALSGNSRSQIKQIVRTITEAGQVDYVEIQESLVPVVERMDRPPCIVCAAWFGKVL 309
Qy      299 IDNMEINL 306
Db      310 IDNIEIHV 317
```

```
RESULT 6
US-10-033-269-4
```

```
; Sequence 4, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Zea mays
US-10-033-269-4
```

```
Query Match      60.4%; Score 959.5; DB 12; Length 323;
Best Local Similarity 62.0%; Pred. No. 2.4e-92;
Matches 191; Conservative 48; Mismatches 62; Indels 7; Gaps 4;

Qy      5 PRVISDKASMSRSRMRAGKLGIVPTMGFLHAGHLSLVAQARQLSD--VVAVSIYV 61
Db      9 PEVIRDKAAMRAWSRRRRRAEGKAVLVPTMGFLHAGHLSLSAASAGPIAVVSIYV 68
Qy      62 NPGQFAPTEDLSTYPSDFDGVKKLASVPGGVVVFHPRNLYDYKNGGGGVDAEA-GGMV 120
Db      69 NPSQFAPTEDLATYPSDFAGDLRKLAAAT-GVAAVFCPPDLYVRGSADRPASAAGSGAV 127
Qy      121 SCVESGSG--HESWVRVEKLELGLCGKSRPVFFRGVATVVTVKLFNIVEPDPVAVFGKKDYQ 178
Db      128 SCLDAGGHAHETWIRVERLEKGLCGSSRPVFFRGVATVVKLFNIVEPDPVAVFGKKDYQ 187
Qy      179 QWRLIQRMVRDLDSIKVIGAEITRDNDGLAMSSRNHLSPEEREKALSINKSLRAKSA 238
Db      188 QWRVICRMVRDLDFAIQVSEVREADGLAMSSRNHLSPEEREKALSIRSLLVDARTA 247
Qy      239 AGDQVHCEKLTNLVIOQVTDAGRIDYAEIVDQNNLEKVEQIKSPVVFCAAWFGKVL 298
Db      248 TLGNSRSQIKQIVRTITEAGQVDYVEIQESLVPVVERMDRPPCIVCAAWFGKVL 307
Qy      299 IDNMEINL 306
Db      308 IDNIEIHV 315
```

```
RESULT 7
US-10-033-269-14
; Sequence 14, Application US/10033269
; Publication No. US20030167503A1
; GENERAL INFORMATION:
; APPLICANT: Gutteridge, Steve
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Genes Encoding Pantothenate Synthetase
; FILE REFERENCE: BB1446 US NA
; CURRENT APPLICATION NUMBER: US/10/033,269
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/247,938
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-033-269-14
```

```
Query Match      58.4%; Score 927.5; DB 12; Length 313;
Best Local Similarity 59.9%; Pred. No. 5.4e-89;
```



```

; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-282
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12221
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12221

Query Match      24.9%; Score 396; DB 15; Length 333;
Best Local Similarity 32.0%; Pred. No. 5.7e-33;
Matches 106; Conservative 50; Mismatches 93; Indels 82; Gaps 9;

QY 30 LVPTMGFLHAGLSLVQAARQLSDV---VAVSIYVNPQCFAPTEBLSYPSDFDGVKKL 86
Db 23 VMTMGALHEGHATLIARTBIAGAEVVTVFVNFPLQFGRGEDLDYPRTLADLKL-- 80
QY 87 ASVPGGVVPHPRNLYDYKNGGDDVAAGMVSVCVSGSHSVRVEKLELG--LCG 144
Db 81 IAEAGADVVPAP-----SADVYPGGFQVRISAGPWRLEG 119
QY 145 KSRVFRFGVATVTKLFNIVEPDVAVFGKKDYQOWRLIQRMVRDLDFSIKIVIGAEITRD 204
Db 120 AFRPGHFDGMLTVVGKLAHLTRPDVALYQKDAQOLALIRWARDLNFGEIVGVPTVRE 179
QY 205 NDGLAMSRNVHLSPEREKALSINKSL-----IRAK-----SA 238
Db 180 DDGLASRNRNYLAADERTALALSQALFAGRDRHAAQEALRARAREVVPATRAAEALSA 239
QY 239 AGDGVHCEKLTNLVIQSVTDAGR-----IDYAEIVQNN 274
Db 240 IGESRAADA---HAAKATPAGTGGPAAVFCARLVLEEARLQPLVLVDYLGVLDPSPD 296
QY 275 LEKV-EQIKSPVFCVAANFGKVLIDNMEI 304
Db 297 FTEIPDFTGFAVLAAVAVGTTRLDNIPL 327

RESULT 11
US-10-113-948-2
; Sequence 2, Application US/10113948
; Publication No. US20030082773A1
; GENERAL INFORMATION:
; APPLICANT: Blundell, Tom L
; APPLICANT: Abell, Christopher
; APPLICANT: von Delft, Frank
; TITLE OF INVENTION: Crystal Structure
; FILE REFERENCE: 620-194
; CURRENT APPLICATION NUMBER: US/10/113,948
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: US 09/659,759
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: PCT/GB01/04067
; PRIOR FILING DATE: 2001-09-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-113-948-2

Query Match      14.5%; Score 231; DB 15; Length 77;
Best Local Similarity 55.3%; Pred. No. 1.5e-16;
Matches 42; Conservative 19; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 148 PVFPRGVATVTVTKLFNIVEPDVAVFGKKDYQOWRLIQRMVRDLDFSIKIVIGAEITRDNDG 207
Db 1 PGHFRGVSTIVSKLFNIVQPDFIACFGKDFQQLALIRKVMADMGFDIEIVGVPIMRKADG 60
QY 208 LAMSSRVNHLSPERE 223
Db 61 LALSSRNGYLTAEQRK 76

RESULT 12
US-09-738-626-6455
; Sequence 6455, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6455
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6455

```

```

Query Match      14.4%; Score 228; DB 10; Length 266;
Best Local Similarity 26.5%; Pred. No. 2e-15;
Matches 82; Conservative 46; Mismatches 117; Indels 64; Gaps 10;

QY 6 RVISDKASMESWSRMRACQKLGIVPTMGFLHAGLSLVQAARQLSDVAVSIYVNPQQ 65
Db 9 RVFTVTEQIRMFSGALRKTKGKPVVLPLGNLGHAGHIALIRAAKRIPGAVVWVAYAGP-- 66
QY 66 FAPTEDLSTYPSDFDGVKKLASVPGGVVV---FHPRNLYDYG---KNGGGDVAAEAGGM 119
Db 67 -----ESDHAELREEL--IDAIFPNFETLWPHGIRVEVTGPTLTPQGA 110
QY 120 VSCVSGSGHESWVRVEKLELGLCGKSRPVFFRGVATVTVTKLFNIVEPDVAVFGKKDYQ 179
Db 111 VTKV-----LGLLG-----ITGATDV-----VLGERDYEL 135
QY 180 WRLLQRMVRDLDFSIKIVIGAEITRDNDGLAMSRNVHLSPEREKALSINKSLRAKSA 239
Db 136 VLVQRALNDLHPTVKLSHVSPTVTRMPDGLAISLRNISVPEDSRRTALSAAALTAGA 195
QY 240 GDGVHCEKLTNLVIQSVTDAGGRIDYAEIVQNNLEKVEQIKSPVFCVAANFGKVLRI 299
Db 196 EHGEAVVKEITVQVLKA---AGVTPDYVEIRGL-DLGAPEIGDARLF-AAITLGDVQLH 250
QY 300 DNMEINLSM 308
Db 251 DNVGLPLGI 259

RESULT 13

```

```
US-10-113-948-1
; Sequence 1, Application US/10113948
; Publication No. US20030082773A1
; GENERAL INFORMATION:
; APPLICANT: Blundell, Tom L
; APPLICANT: Abell, Christopher
; APPLICANT: von Delft, Frank
; TITLE OF INVENTION: Crystal Structure
; FILE REFERENCE: 620-194
; CURRENT APPLICATION NUMBER: US/10/113,948
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: US 09/659,759
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: PCT/GB01/04067
; PRIOR FILING DATE: 2001-09-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-113-948-1

Query Match      11.2%; Score 178; DB 15; Length 79;
Best Local Similarity 51.3%; Pred. No. 5.8e-11;
Matches 39; Conservative 10; Mismatches 25; Indels 2; Gaps 1;

QY 24 OGKLGIVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQGFAPTEDLSTYPSDFDGDV 83
Db 1 EGKRVALVPTMGNLDHGMKLVDEAKARADVIVVNFVPMQFDRPDLARYPRTLQSDC 60

QY 84 KKLASVPGGVVVVFP 99
Db 61 EKLNK--RKVDLVFAP 74

RESULT 14
US-10-128-714-8189
; Sequence 8189, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8189
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8189

Query Match      5.9%; Score 94; DB 15; Length 659;
Best Local Similarity 18.9%; Pred. No. 1;
Matches 66; Conservative 52; Mismatches 107; Indels 124; Gaps 14;

US-10-113-948-1
; Sequence 1, Application US/10113948
; Publication No. US20030082773A1
; GENERAL INFORMATION:
; APPLICANT: Blundell, Tom L
; APPLICANT: Abell, Christopher
; APPLICANT: von Delft, Frank
; TITLE OF INVENTION: Crystal Structure
; FILE REFERENCE: 620-194
; CURRENT APPLICATION NUMBER: US/10/113,948
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: US 09/659,759
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: PCT/GB01/04067
; PRIOR FILING DATE: 2001-09-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-113-948-1

Query Match      11.2%; Score 178; DB 15; Length 79;
Best Local Similarity 51.3%; Pred. No. 5.8e-11;
Matches 39; Conservative 10; Mismatches 25; Indels 2; Gaps 1;

QY 24 OGKLGIVPTMGFLHAGHLSVAQARQLSDVVAVSIYVNPQGFAPTEDLSTYPSDFDGDV 83
Db 1 EGKRVALVPTMGNLDHGMKLVDEAKARADVIVVNFVPMQFDRPDLARYPRTLQSDC 60

QY 84 KKLASVPGGVVVVFP 99
Db 61 EKLNK--RKVDLVFAP 74

RESULT 14
US-10-128-714-8189
; Sequence 8189, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8189
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8189

Query Match      5.9%; Score 94; DB 15; Length 659;
Best Local Similarity 18.9%; Pred. No. 1;
Matches 66; Conservative 52; Mismatches 107; Indels 124; Gaps 14;

US-10-127-032-140
; Sequence 140, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Banguera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 140
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-140

Query Match      5.9%; Score 93.5; DB 15; Length 423;
Best Local Similarity 22.5%; Pred. No. 0.59;
Matches 54; Conservative 31; Mismatches 98; Indels 57; Gaps 9;

QY 70 EDLSTYPSDFDGDVKKLASVPGGVVVVPHRNLYDYGKN---GGGDVABAGGVSCVE 124
Db 50 EQISIPGRDIDEPV--LHHGRGGQTVVHPGNKEFTAGEHIAIRPSGGGGGGGGKAS--N 105

QY 125 SGSGHESWVRVEKLELGLCGKSRPVFPRGVATVVTKLNFIVPEVDVAVFGKKDYQQWRLIQ 184
Db 106 SGEQMDDFV-----FQITQEFLDFMFEDLEPLNVK 137

QY 185 RMVRDLDFSIXVIGAEITRNDGLAMS-----SRNVHLSPEEREKALSINKSLRA 235
Db 138 RHITGTD--TFKTVRAGISNDGNFSRINIVRTLSAHARRIALSGSRAKLRAALKELERI 196

QY 236 KSAAGDGVHCEKLTNLVIQSVTDAGGRIDYAEIVDQNNLEKVEQIKSP-----VVPCV 289
Db 197 KREEDNLGDIQELE----LEIAKLARARIDRVPELDFDLKYNLLVKQPNPTSKAVMFL 252
```

Search completed: December 15, 2003, 13:40:58  
Job time : 32 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:35:45 ; Search time 21 Seconds  
(without alignments)  
624.589 Million cell updates/sec

Title: US-10-033-269-9  
Perfect score: 1588  
Sequence: 1 MAPAPRVISDRASMRWSRS.....AWFGKRLIDNWEINLSMNV 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	35.7	294	4	US-09-107-532A-5153
2	514.5	32.4	283	4	US-09-328-352-7916
3	482.5	30.4	288	4	US-09-134-001C-4321
4	481	30.3	473	4	US-09-252-991A-26805
5	455.5	28.7	279	3	US-09-318-794A-3
6	455.5	28.7	279	3	US-09-318-794A-5
7	103.5	6.5	421	2	US-08-576-626A-53
8	103.5	6.5	1114	2	US-08-576-626A-31
9	100.5	6.3	379	4	US-09-107-532A-4933
10	93.5	5.9	426	4	US-09-252-991A-17427
11	93	5.9	422	4	US-09-679-279-17
12	90.5	5.7	423	4	US-09-328-352-4686
13	90.5	5.7	657	3	US-07-705-490-14
14	90.5	5.7	657	3	US-07-751-891B-14
15	89.5	5.6	373	4	US-09-328-352-7186
16	89	5.6	565	4	US-09-252-991A-16764
17	87.5	5.5	760	4	US-09-252-991A-27790
18	87	5.5	566	4	US-09-252-991A-25622
19	86.5	5.4	1250	3	US-08-938-291A-9
20	86.5	5.4	1250	4	US-09-589-619-9
21	84.5	5.3	223	4	US-09-336-946B-6
22	84.5	5.3	675	4	US-09-252-991A-29071
23	83.5	5.3	2756	1	US-08-375-709-11
24	83.5	5.3	2756	1	US-08-752-929-11
25	83.5	5.3	2756	3	US-09-090-793-7
26	83.5	5.3	2756	4	US-09-231-899-7
27	82.5	5.2	500	4	US-09-328-352-4471

28 82 5.2 497 3 US-08-724-466B-4 Sequence 4, Appli  
29 82 5.2 497 4 US-08-882-164D-4 Sequence 4, Appli  
30 82 5.2 1090 3 US-09-346-237-5 Sequence 5, Appli  
31 81.5 5.1 15281 2 US-08-471-119A-2 Sequence 2, Appli  
32 81 5.1 398 4 US-09-242-859A-4 Sequence 4, Appli  
33 81 5.1 398 4 US-09-242-859A-8 Sequence 8, Appli  
34 81 5.1 436 4 US-09-679-279-4 Sequence 4, Appli  
35 80.5 5.1 236 1 US-08-684-862-5 Sequence 5, Appli  
36 80.5 5.1 930 3 US-09-283-763-2 Sequence 2, Appli  
37 80.5 5.1 930 4 US-09-574-912-2 Sequence 2, Appli  
38 80.5 5.1 1287 4 US-09-252-991A-29606 Sequence 29606, A  
39 79.5 5.0 312 4 US-09-252-991A-17305 Sequence 17305, A  
40 79 5.0 187 3 US-09-095-855-203 Sequence 203, App  
41 79 5.0 187 4 US-09-205-426-203 Sequence 203, App  
42 79 5.0 506 2 US-08-929-501-2 Sequence 2, Appli  
43 79 5.0 506 3 US-09-140-177-2 Sequence 2, Appli  
44 79 5.0 506 3 US-09-397-979-2 Sequence 2, Appli  
45 79 5.0 590 2 US-08-929-501-12 Sequence 12, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-107-532A-5153

; Sequence 5153, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: LYNN A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIORITY DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5153:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 294 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEetical: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...294

; SEQUENCE DESCRIPTION: SEQ ID NO: 5153:

US-09-107-532A-5153

Query Match 35.7%; Score 567; DB 4; Length 294;  
Best Local Similarity 40.7%; Pred. No. 1.9e-55;  
Matches 123; Conservative 51; Mismatches 102; Indels 26; Gaps 4;  
  
QY 6 RVISDKASMRSMRAQCKLIGLVPTMGFLHAGHLSLVAQAQRLSDVVAVSIYVNPQG 65  
DB 17 KTVD---VRSQKWKQKGLVGLVPTMGFLHAGHLSLRKASKENDKVVSIYVNPQTQ 73  
  
QY 66 FAPTEDLSTYPSDFDGDGVKVLASVPGGVVDFVHFRNLVDYKGNKGDDVABAGGMVSCVES 125  
DB 74 FGKNEDLGSPDLERDIE--VCTRGRATAIFNPEVEEMYCDNA----- 115  
  
QY 126 GSGHESWVRVEKLELGLCGKSRPFRGVATVTVTKLFNIVEPDAVFGKDYQOWMLIOR 185  
DB 116 ---STFVNITGTIEGLCGASRPFRHGVCTVTVSKLFNIIIPADRAVFGKDAQQLAVIKR 171  
  
QY 186 MVRDLDFSIKIVIGAEITRDNDGLAMSSRNVLHSPEREKALSKINSLRAKSAAGGVQVH 245  
DB 172 MVRDLNIDVVGPIREEDGLAKSRNTYLSLEERSATILNKSLTLAKELNNGERD 231  
  
QY 246 CEKLTNLVIOQSVTDAG-GRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRLLDNMEI 304  
DB 232 SLKIIETISKNINTYNLAKIDYVEVDLSLQRVNVIEKSVLVAIAVFIKGTLLIDNFTF 291  
  
QY 305 NL 306  
DB 292 EL 293  
  
RESULT 2  
US-09-328-352-7916  
; Sequence 7916, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7916  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7916

Query Match 32.4%; Score 514.5; DB 4; Length 283;  
Best Local Similarity 43.1%; Pred. No. 1.5e-49;  
Matches 122; Conservative 41; Mismatches 97; Indels 23; Gaps 4;  
  
QY 22 RAQCKLIGLVPTMGFLHAGHLSLVAQAQRLSDVVAVSIYVNPQGFAPTEDLSTYPSDFDG 81  
DB 19 RAARKIIGFVPTMGNLHGHLLTVREAKKLCVDDVVVSIFVNPFGQGEDFDNPTLEQ 78  
  
QY 82 DVKKLASVPGGVDFVHFRNLVDYKGNKGDDVABAGGMVSCVESGSGHESWVRVEKLELG 141  
DB 79 DSRLLADV--GCDIIFAPSVEQMYGT-----QPRLLNISVSQITDD 117  
  
QY 142 LCKSRPFRGVATVTVTKLFNIVEPDAVFGKDYQOWMLIORMVRDLDFSIKIVIGAEI 201  
DB 118 LCGSSRPHGDFGVALVTVTKLFNIVQPNYAFQGDYQQLAVIFQFQVDNLPILEVGPVI 177  
  
QY 202 TRNDGLAMSSRNVLHSPEREKALSKINSLRAKSAAGGVQVHCEKLTNLVIOQSVTDAG 261  
DB 178 VRAEDGLALSSRGYLTPEQROVAPVIYQGLKAEQHLHQGKDLQQLVADLK-TLLTDNG 236  
  
QY 262 GRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVRLLDNMEI 304  
DB 237 FVVDYVE-ARQPNLLAASQFDRDVLVFAAKLGGTLLIDNLQV 278

RESULT 3  
US-09-134-001C-4321  
; Sequence 4321, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCI  
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4321  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4321  
  
Query Match 30.4%; Score 482.5; DB 4; Length 288;  
Best Local Similarity 39.1%; Pred. No. 6.3e-46;  
Matches 120; Conservative 45; Mismatches 105; Indels 37; Gaps 6;  
  
QY 6 RVISDKASMRSMRAQCKLIGLVPTMGFLHAGHLSLVAQAQRLSDVVAVSIYVNPQG 65  
DB 5 KVITINEMQSIKQHQREGKTIQFVPTMGALHDGHLTMMKQSVSENDLTVISIFVNPLO 64  
  
QY 66 FAPTEDLSTYPSDFDGD---VKKLASVPGGVDFVHFRNLVDYKGNKGDDVABAGGMVSC 122  
DB 65 FGNEDEDFAYPRQLDDDDAAVKKLQ-----VDYVHPSVDEMYPEELG----- 107  
  
QY 123 VESGSGHESWVRVEKLELGLCGKSRPFRGVATVTVTKLFNIVEPDAVFGKDYQOWML 182  
DB 108 IHLKVGHLAQV-----LEGAQRPGHFEGVTVVWKLNFIVQPDYAFGKKDAQQLAI 159  
  
QY 183 IQRMVRDLDFSIKIVIGAEITRDNDGLAMSSRNVLHSPEREKALSKINSLRAKSAAGDG 242  
DB 160 VEKWKDFNLVPHVIGIDIVREKDGAKSRNIIYLTSEERKEAKHLYQSLRLAKNLYEAG 219  
  
QY 243 QVHCEKLTNLVIOQSVT-----DAGRIDYAEIVDQNNLEKVEQIKSPVFCVAAWFGKVR 297  
DB 220 ----BRDSNEIIQIAAVLNKNISGHIDDLGIYSYPNLIQSKIRHIFISLAVKFSKAR 275  
  
QY 298 LIDNMEI 304  
DB 276 LIDNIII 282  
  
RESULT 4  
US-09-252-991A-26805  
; Sequence 26805, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26805  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26805

```
Query Match      30.3%; Score 481; DB 4; Length 473;
Best Local Similarity 38.7%; Pred. No. 28-45;
Matches 111; Conservative 51; Mismatches 101; Indels 24; Gaps 3;

QY 22 RAQKLIQVPTMGFLHAGHLSVAQARQLSDVAVSIYVNPQCAPTEDLSTYPSDFDG 81
DB 208 RSEGRKIGFVPTMGNLHAGHAAVKKAGRAFPVVIFVNPLOQPSDDLKYPTLAA 267
QY 82 DVKKLASVPGVDVVFHP--RNLVDYKNGKGGDVAEAGGVMSCVSGSGHESWVRVEKLE 139
DB 268 DQERL--LEAGCHLLFTPTVEEMYPDMDG-----QTRIHVPGV 305
QY 140 LGLCGKSRPFRFGVATVTKLNIPEPDVAVFGKDYQOWRLIQMVRDLDFSIKVIGAE 199
DB 306 EGLCGASRPGHFGVATVVKLNVMQPDALFGEKFOQLAVIRKLVRLNLPVQIFGE 365
QY 200 EITRNDGLAMSSRNHLSPEERKALSINKSLRAKSAAGGQVHCEKLTNLVIQSVTD 259
DB 366 PTVRAADGLASRNQYLDQORAAAPAIYRTLRQLGERIRAGAEFPALLADARQALEQ 425
QY 260 AGGRIDYAIIVDONNLEKVEIQIKSPVVFCAAMFGKVRLLIDNMEINL 306
DB 426 AGLRDPYLEIREPISLRPCVPGDQRLVLAAYLGSTRLLIDNLSVHL 472

RESULT 5
US-09-318-794A-3
; Sequence 3, Application US/09318794A
; Patent No. 6177264
; GENERAL INFORMATION:
; APPLICANT: DEGUSSA AKTIENGESILLSCHAFT
; TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC
; FILE REFERENCE: eggeling
; CURRENT APPLICATION NUMBER: US/09/318,794A
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: DE 198 55 312.9
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-318-794A-3
```

```
Query Match      28.7%; Score 455.5; DB 3; Length 279;
Best Local Similarity 41.2%; Pred. No. 6-43;
Matches 121; Conservative 33; Mismatches 93; Indels 47; Gaps 9;

QY 26 KLIGLVPTMGFLHAGHLSVAQARQLSDVAVSIYVNPQCF---APTEDLSTYPSDFDG 82
DB 18 KSVGLVPTMGALHSGHASLVKAARAENDTVVASIFVNPLOQPEALGDCDDYRNPRLQDAD 77
QY 83 VKKLASVPGVDVVFHPNRLYDYGKNGGDVAE--AGGMVSCVSGSGHESWVRVEKLEL 140
DB 78 LALLEE--AGVDIVFAP-----DVEEMPGLPLV-----WARTGSIGT 114
QY 141 GLCGKSRPFRFGVATVTKLNIPEPDVAVFGKDYQOWRLIQMVRDLDFSIKVIGAE 200
DB 115 KLEGASRPGHFGVATVVKLNLVRPDRAYFGQKDAQVAVIRRLVADLDIPVEIRPVP 174
QY 201 ITRNDGLAMSSRNHLSPEERKALSINKSL--LRKSAAGDGQVHCEKLTNLVIQSVT 258
DB 175 IIRGADGLAESNRNQLSADQRAQALVLPQVLSGLQRRKAAGEA-----LDIQAR 225
QY 259 DA-----GGRIDYAIIVDONNLEKVE---QIKSPVVFCAAMFGKVRLLIDNMEI 304
DB 226 DTLASADGVRDLHLEIVDPATLEIDGLLTQPALVVGAI FVGVPVRLIDNIEL 279

RESULT 6
US-09-318-793A-5
```

```
; Sequence 5, Application US/09318793A
; Patent No. 6184007
; GENERAL INFORMATION:
; APPLICANT: Dusch, Nicole
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Puhler, Alfred
; TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC
; FILE REFERENCE: ACID BY ENHANCEMENT OF THE pand GENE IN MICROORGANISMS
; CURRENT APPLICATION NUMBER: US/09/318,793A
; CURRENT FILING DATE: 1999-05-26
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-318-793A-5

Query Match      28.7%; Score 455.5; DB 3; Length 279;
Best Local Similarity 41.2%; Pred. No. 6-43;
Matches 121; Conservative 33; Mismatches 93; Indels 47; Gaps 9;

QY 26 KLIGLVPTMGFLHAGHLSVAQARQLSDVAVSIYVNPQCF---APTEDLSTYPSDFDG 82
DB 18 KSVGLVPTMGALHSGHASLVKAARAENDTVVASIFVNPLOQPEALGDCDDYRNPRLQDAD 77
QY 83 VKKLASVPGVDVVFHPNRLYDYGKNGGDVAE--AGGMVSCVSGSGHESWVRVEKLEL 140
DB 78 LALLEE--AGVDIVFAP-----DVEEMPGLPLV-----WARTGSIGT 114
QY 141 GLCGKSRPFRFGVATVTKLNIPEPDVAVFGKDYQOWRLIQMVRDLDFSIKVIGAE 200
DB 115 KLEGASRPGHFGVATVVKLNLVRPDRAYFGQKDAQVAVIRRLVADLDIPVEIRPVP 174
QY 201 ITRNDGLAMSSRNHLSPEERKALSINKSL--LRKSAAGDGQVHCEKLTNLVIQSVT 258
DB 175 IIRGADGLAESNRNQLSADQRAQALVLPQVLSGLQRRKAAGEA-----LDIQAR 225
QY 259 DA-----GGRIDYAIIVDONNLEKVE---QIKSPVVFCAAMFGKVRLLIDNMEI 304
DB 226 DTLASADGVRDLHLEIVDPATLEIDGLLTQPALVVGAI FVGVPVRLIDNIEL 279

RESULT 7
US-08-576-626A-53
; Sequence 53, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.
; APPLICANT: Staver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,626A
; FILING DATE: 21-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

;/ APPLICATION NUMBER: US-08-576-626A-31  
;/ FILING DATE: 21-DEC-1995  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Dianne Casuto  
;/ REGISTRATION NUMBER: P-40,943  
;/ REFERENCE/DOCKET NUMBER: 5857.US.01  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (847) 938-3137  
;/ TELEFAX: (847) 938-2623  
;/ TELEX:  
;/ INFORMATION FOR SEQ ID NO: 53:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 421 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: No. 5998194e  
US-08-576-626A-53

Query Match 6.5%; Score 103.5; DB 2; Length 421;  
Best Local Similarity 21.6%; Pred. No. 0.0064;  
Matches 74; Conservative 39; Mismatches 109; Indels 121; Gaps 17;  
QY 14 MRGWSRMRQAQGLIGLVP-TMGFLHAGHLSVAQAQRLSDVAVSVIYVNPQGFAPTEDL 72  
Db 1 MRVVFSSMASKSHLFGVLPLAWAFRAAGH-----EVRVVASP-----ALTEDI 43  
QY 73 STYPSDFDGDVKKLASVPGGVDVVFPHRNLYDYKNGGGDVAEAGGMVSCVSGSGHESW 132  
Db 44 T-----AAGLTAVPVGTDV-----DLVDFMTHAGHDIIIDYVRSLSDFSERDPATLTW 89  
QY 133 VRVEKLE-----LGLCGKSRP-----VFRGVATVVTVKLFN- 163  
Db 90 EHLRGQTVLTPTTFYALMSPDTLIEGMVSFCRKWRPDLVIWEPLTFAAPIAGAVTGTGPHA 149  
QY 164 --IIVPEDVAVFGKKDY-----LGLCGKSRP-----VFRGVATVVTVKLFN- 182  
Db 150 RLILWGFDTITRAQNFGLLPPDQPEHREGPLAEWLTWLEKYGGPAFDEEVVVGQWTID 209  
QY 183 -IQRMVRDLDFSIVKIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGD 241  
Db 210 PAPAAR-LDTGLKTGMRYV-DYNGPSVVPWMLHDEPERRRVCLTLGTS-SRENSI--- 263  
QY 242 GOVHCEKLTNLVIOQSVTDAGGRIDYAEIV---DONNLEKVEQI 281  
Db 264 GOVSIEELGAV-----GDVD-AEIIATFDAQOLEGVANI 297

RESULT 8  
US-08-576-626A-31  
; Sequence 31, Application US/08576626A  
; Patent No. 5998194  
; GENERAL INFORMATION:  
; APPLICANT: Summers, R.G.  
; APPLICANT: Katz, L.  
; APPLICANT: Donadio, S.  
; APPLICANT: Staver, M.J.  
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR  
; TITLE OF INVENTION: BIOSYNTHESIS GENES  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:

;/ APPLICATION NUMBER: US/08/576,626A  
;/ FILING DATE: 21-DEC-1995  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER:  
;/ FILING DATE:  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Dianne Casuto  
;/ REGISTRATION NUMBER: P-40,943  
;/ REFERENCE/DOCKET NUMBER: 5857.US.01  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (847) 938-3137  
;/ TELEFAX: (847) 938-2623  
;/ TELEX:  
;/ INFORMATION FOR SEQ ID NO: 31:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 1114 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: No. 5998194e  
US-08-576-626A-31

Query Match 6.5%; Score 103.5; DB 2; Length 1114;  
Best Local Similarity 21.6%; Pred. No. 0.03;  
Matches 74; Conservative 39; Mismatches 109; Indels 121; Gaps 17;  
QY 14 MRSWSRMRQAQGLIGLVP-TMGFLHAGHLSVAQAQRLSDVAVSVIYVNPQGFAPTEDL 72  
Db 361 MRVVFSSMASKSHLFGVLPLAWAFRAAGH-----EVRVVASP-----ALTEDI 403  
QY 73 STYPSDFDGDVKKLASVPGGVDVVFPHRNLYDYKNGGGDVAEAGGMVSCVSGSGHESW 132  
Db 404 T-----AAGLTAVPVGTDV-----DLVDFMTHAGHDIIIDYVRSLSDFSERDPATLTW 449  
QY 133 VRVEKLE-----LGLCGKSRP-----VFRGVATVVTVKLFN- 163  
Db 450 EHLRGQTVLTPTTFYALMSPDTLIEGMVSFCRKWRPDLVIWEPLTFAAPIAGAVTGTGPHA 509  
QY 164 --IIVPEDVAVFGKKDY-----LGLCGKSRP-----VFRGVATVVTVKLFN- 182  
Db 510 RLILWGFDTITRAQNFGLLPPDQPEHREGPLAEWLTWLEKYGGPAFDEEVVVGQWTID 569  
QY 183 -IQRMVRDLDFSIVKIGAEITRDNDGLAMSRNVHLSPEEREKALSINKSLRAKSAAGD 241  
Db 570 PAPAAR-LDTGLKTGMRYV-DYNGPSVVPWMLHDEPERRRVCLTLGTS-SRENSI--- 623  
QY 242 GOVHCEKLTNLVIOQSVTDAGGRIDYAEIV---DONNLEKVEQI 281  
Db 624 GOVSIEELGAV-----GDVD-AEIIATFDAQOLEGVANI 657

RESULT 9  
US-09-107-532A-4933  
; Sequence 4933, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 4933:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...379
; SEQUENCE DESCRIPTION: SEQ ID NO: 4933:
US-09-107-532A-4933

Query Match      6.3%; Score 100.5; DB 4; Length 379;
Best Local Similarity 23.3%; Pred. No. 0.012;
Matches 74; Conservative 34; Mismatches 97; Indels 113; Gaps 17;

QY 50 QLSGV-VAVSIYVNPQGFAPTELDSTYPSDFDGVKKLASVPGGVV--VFHP----- 99
Db 61 EMSDLAIKTCAYLS--QLLKVEDAQIVSSASAGIAQSVAAIIGQDMVHYHPYTEKIAR 118
QY 100 RNL-----YDVG-----KNGGGVBAAGMVCVSGSGHESVVRVEKLEGLCGK 145
Db 119 REIIPKGNVDYGTPEVVMVAQGGKVIBAGYANMCTPE--HVSMMITEKTAALLYIK 175
QY 146 SRPFVRGVATV-----VTK-----LPNIVEPDV----- 169
Db 176 SHAVOKSMUTVAVMEVVAKAHLPLIVDAAAEEDLFKYSSEMGVDLVIYSGAKAIEGPSA 235
QY 170 -AVFGKKDYQOQWRLIQRWDRDLDFSIKVIQ--AEITRDNDGLAMSSRNVLHSPEREKAL 236
Db 236 GLVIGKKYIQWIRLQS-----KGIGRAMKIGKN-----IL 267
QY 227 SINKSL---LRKSAAGDGVH-----CEKLTNLVIOQSVTDAGRIDYAEIV-----D 271
Db 268 GFTQALEDLKNGSESGSKMKQLAPPVNAINQDPLSAKIVQDGAGRIDYRASVTINGD 327
QY 272 QNNLEKVEQIK--SPVVE 287
Db 328 KSAKEVIQELKARNPAVI 345

RESULT 10
US-09-252-991A-17427
; Sequence 17427, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
```

```
;
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17427
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17427

Query Match      5.9%; Score 93.5; DB 4; Length 426;
Best Local Similarity 22.5%; Pred. No. 0.088;
Matches 54; Conservative 31; Mismatches 98; Indels 57; Gaps 9;

QY 70 EDLSTYPSDFDGVKKLASVPGGVVVFHPRNL-YDYGN-----GGDVAAEAGMVCV 124
Db 53 EQISIPERDIDEV--LHHGRGGRQTVVHPGNKEFTAGBHIARPSGGGGGRGGKAS--N 108
QY 125 SGGSHESVVRVEKLEGLCGKSRPVFFRGVATVVTVKLENIVPDVAVFGKDYQOQWRLIQ 184
Db 109 SGEQMDDFV-----FQITQEFDFPFDELELPNLVK 140
QY 185 RMVRLDIFSIVKIGABITRDNDGLAMS-----SRNVHLSPEEREKALSINKSILRA 235
Db 141 RHITGTD-TFKTVRAGISNDGNPSRINIVITLRSAAHRIALSGGSRAKLRAALKELERI 199
QY 236 KSAAGDGVHCEKLTNLVIOQSVTDAGRIDYAEIVDQNNLEKVEQIKSP-----YVFCV 289
Db 200 KREPDNLGDIQELE---LEIAKLRARIDRVFPDFTDLKYNLLVKPNPTSKAVNFCL 255

RESULT 11
US-09-679-279-17
; Sequence 17, Application US/09679279
; Patent No. 6524841
; GENERAL INFORMATION:
; APPLICANT: McDaniel, Robert
; APPLICANT: Volchegursky, Yanina
; TITLE OF INVENTION: Recombinant Megalomicin Biosynthetic
; FILE REFERENCE: 300622004700
; CURRENT APPLICATION NUMBER: US/09/679,279
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/158,305
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 60/190,024
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Micromonospora megalomicea
US-09-679-279-17

Query Match      5.9%; Score 93; DB 4; Length 422;
Best Local Similarity 24.2%; Pred. No. 0.099;
Matches 64; Conservative 32; Mismatches 99; Indels 70; Gaps 14;

QY 14 MRSWRSRMAQKGLIGLVP-TWGFHAGHLSLVAQARQLSDVVAVIYVNPQGFAPTEDL 72
Db 1 MRVVFSSMAKSHLFGVLPLAWAFRAAGH-----EVRVVASP---ALTDI 43
QY 73 STYPSDFDGVKKLASVPGGVVVFHPRNLVDYKNGGGDVAAEAGMVCVSGSGHESW 132
Db 44 T-----AAGLTAVPVGTIV-----DLVDFMTHAGYDIIDYVRSLDPSRDPATSTW 89
QY 133 VRVEKLEGLCGKSRPVFFRGVATVVTVKLF-NIVPDVAVFGKDY-QOQWRLIQRWRL 190
Db 90 DHL-----LGMQTVLTPTFYALMSPDSLVGWMISFCRSWR-----PDM 127
QY 191 DFSIKVIGAEITRDNDGLAMSSR--NVHLSPEEREKALSINKSILRAKSAAGDGVHCE 247
Db 128 SSGPQTFASIAATVTGVAHARLLMGPDITVRAROKFL---GLLPQFPAA-----HRED 178
```

```

QY 248 KLTNLVIOQVTDAGRI--DYAEIV 270
Db 179 PLAEWLTWVERGGRVPODVEELV 203

RESULT 12
US-09-328-352-4686
; Sequence 4686, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4686
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4686

Query Match 5.7%; Score 90.5; DB 4; Length 423;
Best Local Similarity 22.3%; Pred. No. 0.19; Indels 115; Gaps 20;
Matches 69; Conservative 43; Mismatches 82; Indels 115; Gaps 20;

QY 8 ISDKASMRWSRSMRAQGLIGLIVPTMGFLHAGHLSLVAQARQLSDVAVSIY-----V 61
Db 145 LSPNAS--GYAEQVLAQSLIFEVFN-----GLDADSAAWTEPMAVHALHAYRSRV 194
QY 62 NPGQFAPTEDELSTYP-----SDFDGVKKLASVPGGVGVVFPRLM 102
Db 195 KTGE--PATVIGCGPVGLGVILMLKAAAGVKTWVASDFSPNRRLAB--QCGADIVVDPKET 251
QY 103 YDYG--KNGG--GDVAEAGGWSVCVBSGSG-----HESWVRVEKLELGLCGKS 146
Db 252 SFANWKEFGLLONVSDA-----INMGMLFDKIQATLPWWHGWRMTDK--LGALPK- 302
QY 147 RVFFRGVATVTKLFNIVEPDAVFGKDKYQQWRLIQRMVRDLDFSIKIVGAETRDND 206
Db 303 RVIFPCVGV-----PGV-----LQIIEGAPLFSIVGV----- 332
QY 207 GLAMSRNVHLSPEEREKALSINKSL-----LRKSA---AGDQVHCEKL--TNL 252
Db 333 GYCMQSDKI-----EPALAINKELEIQVLGYTPLEFRDALHMAEGKNCSPITGV 385
QY 253 V-IQSVTDA 260
Db 386 VQLEGVTNA 394

RESULT 13
US-07-705-490-14
; Sequence 14, Application US/07705490
; Patent No. 6107025
; GENERAL INFORMATION:
; APPLICANT: Caskey, C. T.
; APPLICANT: Nelson, David L.
; APPLICANT: Pieretti, Maura
; APPLICANT: Warren, Stephen T.
; APPLICANT: Costra, Ben A.
; TITLE OF INVENTION: Diagnosis of the Fragile X Syndrome
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas D. Paul
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/705,490
; FILING DATE: 19910708
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5350
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/651-5325
; TELEFAX: 713/651-5246
; TELEX: 762829
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 657 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: C-terminal
US-07-705-490-14

Query Match 5.7%; Score 90.5; DB 3; Length 657;
Best Local Similarity 19.5%; Pred. No. 0.38; Indels 127; Gaps 11;
Matches 61; Conservative 34; Mismatches 91; Indels 127; Gaps 11;

QY 34 MGFLHAGHLSLVAQARQLSDVAVSIYVNPQFAPTEDELST---YPSDFDGDVKKLASVP 90
Db 295 MGLAIGTGANIQAQKVPVGTADL-----DEDTCTFIYGEDQDAVKKARSFLE 345
QY 91 GGVVVFPRLNLYDVGKNGG--GDVAEAGGWSVC----- 122
Db 346 FAEDVIQVPRNLVIVGKNGKLIQEIYDKSGVVVRVRIEAEKNVPEEEIIMPNSLPNN 405
QY 123 -----VESGSGHESWVRVEKLELGLCGKSRVFFRGV-----ATVVTKL 161
Db 406 SRVGNPAPEEKHLDIKENSTHFSQPNSTKVQRGV-----PFVFGTKDSANATVLL-- 459
QY 162 FNIVEPDAVFGKDYQQWRLIQRMVRDLDFSIKIVGAE-----ITRDNDG 207
Db 460 -----DYHLNLYKEVDQLRL--ERL--QIDQLRQIGASSRPPPNRTDKESYVTDGQG 510
QY 208 LAMSSR-----NVHLSPEEREKALSINK 230
Db 511 MGRGSRPYRNRHGRRGPGYTSNTNSEASNASSETSDHRDELSDWSLAPTEERESEFLRR 570
QY 231 SLLRAKSAAGDQG 243
Db 571 GDGRRRGGRGQ 583

RESULT 14
US-07-751-891B-14
; Sequence 14, Application US/07751891B
; Patent No. 6180337
; GENERAL INFORMATION:
; APPLICANT: Caskey, C. T.
; APPLICANT: Nelson, David L.
; APPLICANT: Pieretti, Maura
; APPLICANT: Warren, Stephen T.
; APPLICANT: Costra, Ben A.
; APPLICANT: Fu, Ying-hui
; TITLE OF INVENTION: Diagnosis of the Fragile X Syndrome
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas D. Paul
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.

```

